

A;Note: the authors translated the codon ATG for residues 63, 66 and 239 as Trp and GGA for residue 239 as Glu

R;Stamenkovic, I.; Amiot, M.; Pesando, J.M.; Seed, B.

Cell 56, 1057-1062, 1989

A;Title: A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family.

A;Reference number: A32376; MUID:89168434; PMID:2466575

A;Accession: A32376

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-238,'E',240-361 <STA>

A;Cross-references: GB:M24915; NID:g180196; PIDN:AAA35674.1; PID:g180197

R;Bosch, P.P.; Stevens, J.W.; Buckwalter, J.A.; Midura, R.J.

submitted to the EMBL Data Library, November 1995

A;Reference number: H00921

A;Accession: G02251

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-25,'M',27-108,'S',110-361 <BOS>

A;Cross-references: EMBL:U40373; NID:g1101785; PID:g1101786

R;Goldstein, L.A.; Zhou, D.F.H.; Picker, L.J.; Minty, C.N.; Bargatze, R.F.;

Ding, J.F.; Butcher, E.C.

Cell 56, 1063-1072, 1989

A;Title: A human lymphocyte homing receptor, the hermes antigen, is related to cartilage proteoglycan core and link proteins.

A;Reference number: A32377; MUID:89168435; PMID:2466576

A;Accession: A32377

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-108,'S',110-293,'S' <GOL>

A;Cross-references: GB:M25078; NID:g186660; PIDN:AAA36138.1; PID:g186661

C;Superfamily: human cell adhesion protein CD44

C;Keywords: alternative splicing; cell adhesion; surface antigen; transmembrane protein

F;269-285/Domain: transmembrane #status predicted <TMM>

Query Match 82.4%; Score 28; DB 2; Length 361;

Best Local Similarity 83.3%; Pred. No. 65;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFST 6

|| |||

Db 195 FYTFST 200

RESULT 13

I77371

CD44R5 - human

C;Species: Homo sapiens (man)

C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000

C;Accession: I77371

R;Tanabe, K.K.; Nishi, T.; Saya, H.

Mol. Carcinog. 7, 212-220, 1993

A;Title: Novel variants of CD44 arising from alternative splicing: changes in the CD44 alternative splicing pattern of MCF-7 breast carcinoma cells treated with hyaluronidase.

A;Reference number: I57483; MUID:93356912; PMID:8352881

A;Accession: I77371
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-395 <RES>
A;Cross-references: GB:S66400; NID:g435697; PIDN:AAB27919.1; PID:g435700
C;Genetics:
A;Gene: GDB:CD44
A;Cross-references: GDB:120739; OMIM:107269
A;Map position: 11pter-11p13
A;Introns: 257/1
C;Superfamily: human cell adhesion protein CD44

Query Match 82.4%; Score 28; DB 2; Length 395;
Best Local Similarity 83.3%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFST 6
|| |||
Db 195 FYTFST 200

RESULT 14

JH0518

lymphocyte homing receptor CD44, splice form CD44R1 - human
N;Alternate names: cell adhesion molecule core protein CD44E, keratinocyte; cell surface glycoprotein CD44

N;Contains: lymphocyte homing receptor CD44, splice form CD44R1; lymphocyte homing receptor CD44, splice form CD44R2

C;Species: Homo sapiens (man)

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C;Accession: JH0518; JH0519; PH0859; A39209; A42402; C42402; A53029; S16147

R;Dougherty, G.J.; Lansdorp, P.M.; Cooper, D.L.; Humphries, R.K.

J. Exp. Med. 174, 1-5, 1991

A;Title: Molecular cloning of CD44R1 and CD44R2, two novel isoforms of the human CD44 lymphocyte "homing" receptor expressed by hemopoietic cells.

A;Reference number: JH0518; MUID:91277598; PMID:2056274

A;Accession: JH0518

A;Molecule type: mRNA

A;Residues: 1-426 <DOU>

A;Cross-references: UNIPROT:Q9UCB0

A;Experimental source: lymphocytes, cell line KG1a

A;Accession: JH0519

A;Molecule type: mRNA

A;Residues: 1-223,288-426 <DO2>

A;Experimental source: lymphocyte, cell line KG1a

R;Cooper, D.L.; Dougherty, G.; Harn, H.J.; Jackson, S.; Baptist, E.W.; Byers, J.; Datta, A.; Phillips, G.; Isola, N.R.

Biochem. Biophys. Res. Commun. 182, 569-578, 1992

A;Title: The complex CD44 transcriptional unit: alternative splicing of three internal exons generates the epithelial form of CD44.

A;Reference number: PH0859; MUID:92134271; PMID:1734871

A;Accession: PH0859

A;Molecule type: DNA

A;Residues: 223-357 <COO>

R;Brown, T.A.; Bouchard, T.; St. John, T.; Wayner, E.; Carter, W.G.

J. Cell Biol. 113, 207-221, 1991

A;Title: Human keratinocytes express a new CD44 core protein (CD44E) as a
 heparan-sulfate intrinsic membrane proteoglycan with additional exons.
 A;Reference number: A39209; MUID:91177958; PMID:2007624
 A;Accession: A39209
 A;Molecule type: mRNA
 A;Residues: 184-376 <BRO>
 A;Cross-references: GB:X55938; NID:g29802; PIDN:CAA39404.1; PID:g930047
 R;Jackson, D.G.; Buckley, J.; Bell, J.I.
 J. Biol. Chem. 267, 4732-4739, 1992
 A;Title: Multiple variants of the human lymphocyte homing receptor CD44
 generated by insertions at a single site in the extracellular domain.
 A;Reference number: A42402; MUID:92165834; PMID:1537855
 A;Accession: A42402
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 217-223,288-359 <JAC>
 A;Note: sequence extracted from NCBI backbone (NCBIN:83964, NCBIP:83965)
 A;Note: variant B
 A;Accession: C42402
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 217-320 <JA2>
 A;Note: sequence extracted from NCBI backbone (NCBIN:83968, NCBIP:83969)
 A;Note: variant D
 R;Shepley, M.P.; Racaniello, V.R.
 J. Virol. 68, 1301-1308, 1994
 A;Title: A monoclonal antibody that blocks poliovirus attachment recognizes the
 lymphocyte homing receptor CD44.
 A;Reference number: A53029; MUID:94149816; PMID:7508992
 A;Accession: A53029
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 67-76, 'X', 78-89 <SHE>
 C;Genetics:
 A;Gene: GDB:CD44; MDU2; MDU3; MI
 A;Cross-references: GDB:120739; OMIM:107269
 A;Map position: 11pter-11p13
 A;Introns: 35/1; 65/1; 133/1
 C;Superfamily: human cell adhesion protein CD44
 C;Keywords: alternative splicing; cell adhesion; chondroitin sulfate
 proteoglycan; glycoprotein
 F;1-426/Product: lymphocyte homing receptor CD44, splice form CD44R1 #status
 predicted <MA1>
 F;1-223,288-426/Product: lymphocyte homing receptor CD44, splice form CD44R2
 #status predicted <MA2>
 F;299/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;354/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 82.4%; Score 28; DB 2; Length 426;
 Best Local Similarity 83.3%; Pred. No. 76;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYXFST 6
 || |||
 Db 195 FYTFST 200

RESULT 15

A47442

olfactomedin precursor - bullfrog

C;Species: Rana catesbeiana (bullfrog)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: A47442

R;Yokoe, H.; Anholt, R.R.H.

Proc. Natl. Acad. Sci. U.S.A. 90, 4655-4659, 1993

A;Title: Molecular cloning of olfactomedin, an extracellular matrix protein specific to olfactory neuroepithelium.

A;Reference number: A47442; MUID:93281637; PMID:8506313

A;Accession: A47442

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-464 <YOK>

A;Cross-references: UNIPROT:Q07081; GB:L13595; NID:g294501; PIDN:AAA49527.1;

PID:g294502

C;Keywords: extracellular matrix

Query Match 82.4%; Score 28; DB 2; Length 464;
Best Local Similarity 71.4%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
|||
Db 411 FYMFDTK 417

Search completed: February 10, 2005, 15:59:25

Job time : 12.8451 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 51.0704 Seconds
(without alignments)
70.188 Million cell updates/sec

Title: US-10-067-484-4
Perfect score: 34
Sequence: 1 FYXFSTK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*
 1: uniprot_sprot:*
 2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	33	97.1	41	2	Q91FZ4	Q91fz4 chilo iride
2	32	94.1	108	2	Q6EB60	Q6eb60 campylobact
3	32	94.1	615	2	Q8IJP0	Q8ijp0 plasmodium
4	30	88.2	60	2	Q8F8P3	Q8f8p3 leptospira
5	30	88.2	116	1	YEB0_YEAST	P40000 saccharomyc
6	30	88.2	199	2	Q87NM8	Q87nm8 vibrio para
7	30	88.2	201	1	Y787_HAEIN	P44052 haemophilus
8	30	88.2	477	2	Q642R1	Q642r1 xenopus lae
9	29	85.3	94	1	YBFE_BACSU	O31445 bacillus su
10	29	85.3	100	2	Q65XA7	Q65xa7 oryza sativ
11	29	85.3	202	2	Q6LPX2	Q6lpx2 photobacter
12	29	85.3	207	2	Q63BG4	Q63bg4 bacillus ce
13	29	85.3	209	2	Q81QL8	Q81ql8 bacillus an
14	29	85.3	209	2	Q6HIX3	Q6hix3 bacillus th
15	29	85.3	222	2	Q9YMQ2	Q9ymq2 lymantria d
16	29	85.3	224	2	Q8VSW2	Q8vsw2 staphylococ
17	29	85.3	227	2	Q6SX64	Q6sx64 human cytom
18	29	85.3	227	2	Q6SXB5	Q6sxb5 human cytom
19	29	85.3	258	1	UPPS_THEAC	Q9hkhq0 thermoplasm
20	29	85.3	258	1	UPPS_THEVO	Q97b58 thermoplasm
21	29	85.3	258	2	Q9M2I5	Q9m2i5 arabidopsis
22	29	85.3	258	2	Q6MEF4	Q6mef4 parachlamyd
23	29	85.3	262	2	Q863K2	Q863k2 sus scrofa
24	29	85.3	297	2	Q8AV24	Q8av24 fugu rubrip
25	29	85.3	348	2	Q75D24	Q75d24 ashbya goss
26	29	85.3	369	2	Q73VJ4	Q73vj4 mycobacteri
27	29	85.3	370	1	DDL_MYCBO	Q7txh9 mycobacteri
28	29	85.3	373	1	DDL_MYCSM	Q9zgn0 mycobacteri
29	29	85.3	373	1	DDL_MYCTU	P95114 mycobacteri
30	29	85.3	373	2	Q18197	Q18197 caenorhabdi
31	29	85.3	384	1	DDL_MYCLE	Q9cbs0 mycobacteri
32	29	85.3	491	1	ZAPA_PROMI	Q11137 proteus mir
33	29	85.3	491	2	O85374	O85374 proteus mir
34	29	85.3	555	2	Q64QR5	Q64qr5 bacteroides
35	29	85.3	635	2	Q9VUK7	Q9vuk7 drosophila
36	29	85.3	700	2	Q720Z1	Q720z1 listeria mo
37	29	85.3	1435	2	Q7S9Z4	Q7s9z4 neurospora
38	29	85.3	1976	2	Q7RTC8	Q7rtc8 plasmodium
39	28	82.4	84	2	Q7YT13	Q7yt13 rhodnius pr
40	28	82.4	103	2	Q6ZI03	Q6zi03 oryza sativ
41	28	82.4	148	2	Q6ME93	Q6me93 parachlamyd
42	28	82.4	207	2	Q8UP61	Q8up61 human immun

43	28	82.4	224	2	Q63IP1	Q63ip1 burkholderi
44	28	82.4	229	2	Q8VS70	Q8vs70 borrelia he
45	28	82.4	230	1	UPPS_BORBU	O51146 borrelia bu

ALIGNMENTS

RESULT 1

Q91FZ4

ID Q91FZ4 PRELIMINARY; PRT; 41 AA.
AC Q91FZ4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 164R.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99125223; PubMed=9926400; DOI=10.1023/A:1008017820941;
RA Muller K., Tidona C.A., Bahr U., Darai G.;
RT "Identification of a thymidylate synthase gene within the genome of
RT Chilo iridescent virus.";
RL Virus Genes 17:243-258(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93118242; PubMed=1475907;
RA Sonntag K.C., Darai G.;
RT "Characterization of the third origin of DNA replication of the genome
RT of insect iridescent virus type 6.";
RL Virus Genes 6:333-342(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94353641; PubMed=8073636;
RA Sonntag K.C., Schnitzler P., Koonin E.V., Darai G.;
RT "Chilo iridescent virus encodes a putative helicase belonging to a
RT distinct family within the 'DEAD/H' superfamily: implications for the
RT evolution of large DNA viruses.";
RL Virus Genes 8:151-158(1994).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=94292906; PubMed=8021587;
RA Schnitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
RA Koonin E.V., Darai G.;
RT "Insect iridescent virus type 6 encodes a polypeptide related to the
RT largest subunit of eukaryotic RNA polymerase II.";
RL J. Gen. Virol. 75:1557-1567(1994).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213160; PubMed=7698884;
RA Sonntag K.C., Schnitzler P., Janssen W., Darai G.;
RT "Identification of the primary structure and the coding capacity of
RT the genome of insect iridescent virus type 6 between the genome
RT coordinates 0.310 and 0.347 (7990 bp).";
RL Intervirology 37:287-297(1994).

RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98141693; PubMed=9482589; DOI=10.1023/A:1007932620930;
 RA Bahr U., Tidona C.A., Darai G.;
 RT "The DNA sequence of Chilo iridescent virus between the genome
 RT coordinates 0.101 and 0.391; similarities in coding strategy between
 RT insect and vertebrate iridoviruses.";
 RL Virus Genes 15:235-245(1997).
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Delius H., Darai G., Fluegel R.M.;
 RT "DNA analysis of insect iridescent virus 6: evidence for circular
 RT permutation and terminal redundancy.";
 RL J. Virol. 49:609-614(1984).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86174607; PubMed=3959991;
 RA Lorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
 RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
 RT in mice.";
 RL Med. Microbiol. Immunol. 175:43-53(1986).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87321126; PubMed=2820141;
 RA Schnitzler P., Soltau J.B., Fischer M., Reisner H., Scholz J.,
 RA Delius H., Darai G.;
 RT "Molecular cloning and physical mapping of the genome of insect
 RT iridescent virus type 6: further evidence for circular permutation of
 RT the viral genome.";
 RL Virology 160:66-74(1987).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89073752; PubMed=3201750;
 RA Fischer M., Schnitzler P., Delius H., Darai G.;
 RT "Identification and characterization of the repetitive DNA element in
 RT the genome of insect iridescent virus type 6.";
 RL Virology 167:485-496(1988).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92196996; PubMed=1549908;
 RA Handermann M., Schnitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
 RA Darai G.;
 RT "Identification and mapping of origins of DNA replication within the
 RT DNA sequences of the genome of insect iridescent virus type 6.";
 RL Virus Genes 6:19-32(1992).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93260401; PubMed=8492091;
 RA Stohwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;
 RT "Identification of the gene encoding the major capsid protein of
 RT insect iridescent virus type 6 by polymerase chain reaction.";
 RL J. Gen. Virol. 74:873-879(1993).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94167241; PubMed=8121799;
 RA Schnitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
 RA Delius H., Darai C.;

RT "Identification of genes encoding zinc finger proteins, non-histone
 RT chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
 RT in the genome of Chilo iridescent virus.";
 RL Nucleic Acids Res. 22:158-166(1994).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99383793; PubMed=10456793; DOI=10.1023/A:1008072319875;
 RA Muller K., Tidona C.A., Darai G.;
 RT "Identification of a gene cluster within the genome of Chilo
 RT iridescent virus encoding enzymes involved in viral DNA replication
 RT and processing.";
 RL Virus Genes 18:243-264(1999).
 RN [15]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21342589; PubMed=11448171; DOI=10.1006/viro.2001.0963;
 RA Jakob N.J., Muller K., Bahr U., Darai G.;
 RT "Analysis of the first complete DNA sequence of an invertebrate
 RT iridovirus: coding strategy of the genome of Chilo iridescent virus.";
 RL Virology 286:182-196(2001).
 DR EMBL; AF303741; AAK82038.1; -.
 SQ SEQUENCE 41 AA; 4830 MW; 015CE2869B5DEBAE CRC64;

Query Match 97.1%; Score 33; DB 2; Length 41;
 Best Local Similarity 85.7%; Pred. No. 2.2;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
 || ||||
 Db 6 FYSFSTK 12

RESULT 2

Q6EB60

ID Q6EB60 PRELIMINARY; PRT; 108 AA.
 AC Q6EB60;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Tgh072 (Fragment).
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TGH 9011;
 RX PubMed=15231810; DOI=10.1128/JB.186.14.4781-4795.2004;
 RA Poly F., Threadgill D., Stintzi A.;
 RT "Identification of Campylobacter jejuni ATCC 43431-Specific Genes by
 RT Whole Microbial Genome Comparisons.";
 RL J. Bacteriol. 186:4781-4795(2004).
 DR EMBL; AY501952; AAS99025.1; -.
 FT NON_TER 108 108
 SQ SEQUENCE 108 AA; 13039 MW; 8FF92B4F99711A5D CRC64;

Query Match 94.1%; Score 32; DB 2; Length 108;
 Best Local Similarity 85.7%; Pred. No. 10;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
|| |||
Db 12 FYLFSTK 18

RESULT 3

Q8IJPO

ID Q8IJPO PRELIMINARY; PRT; 615 AA.
AC Q8IJPO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF10_0152;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014831; AAN35350.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001201; PAP_25A_core.
DR InterPro; IPR002058; PAP_assoc.
KW Hypothetical protein.
SQ SEQUENCE 615 AA; 72618 MW; FD6BEBF3F5D46C33 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 615;
Best Local Similarity 85.7%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
|| |||
Db 29 FYEFSTK 35

RESULT 4

Q8F8P3

ID Q8F8P3 PRELIMINARY; PRT; 60 AA.
AC Q8F8P3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Hypothetical protein.
 GN OrderedLocusNames=LA0513;
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
 RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
 RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
 RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
 RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
 RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
 RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
 RA Xu J.-G., Zhao G.-P.;
 RT "Unique physiological and pathogenic features of Leptospira
 RT interrogans revealed by whole-genome sequencing.";
 RL Nature 422:888-893(2003).
 DR EMBL; AE011237; AAN47711.1; -.
 KW Complete proteome.
 SQ SEQUENCE 60 AA; 7216 MW; 5B4F327D42EDE78E CRC64;

Query Match 88.2%; Score 30; DB 2; Length 60;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYXFSTK 7
 |||:
 D5 21 FYSFATK 27

RESULT 5

YEB0_YEAST

ID YEB0_YEAST STANDARD; PRT; 116 AA.
 AC P40000;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Hypothetical 13.5 kDa protein in GLC3-GCN4 intergenic region.
 GN OrderedLocusNames=YEL010W;
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=97313264; PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
 RA Hunicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";

RL Nature 387:78-81(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RA Marsischky G., Rolfs A., Richardson A., Kane M., Baqui M., Taycher E.,
 RA Hu Y., Vannberg F., Weger J., Kramer J., Moreira D., Kelley F.,
 RA Zuo D., Raphael J., Hogle C., Jepson D., Williamson J., Camargo A.,
 RA Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E.,
 RA LaBaer J.;
 RT "Creation of the YFLEX clone resource: cloning of *Saccharomyces*
 RT *cerevisiae* ORFs in the Gateway recombinational cloning system.";
 RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U18530; AAB64487.1; -.
 DR EMBL; AY558340; AAS56666.1; -.
 DR PIR; S50449; S50449.
 DR GermOnline; 139014; -.
 DR SGD; S000000736; YEL010W.
 KW Hypothetical protein..
 SQ SEQUENCE 116 AA; 13523 MW; 66B8654F75C708AC CRC64;

Query Match 88.2%; Score 30; DB 1; Length 116;
 Best Local Similarity 71.4%; Pred. No. 33;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYXFSTK 7
 :| |||
 Db 31 YYSFSTK 37

RESULT 6

Q87NM8

ID Q87NM8 PRELIMINARY; PRT; 199 AA.
 AC Q87NM8;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein VP1840.
 GN OrderedLocusNames=VP1840;
 OS *Vibrio parahaemolyticus*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; *Vibrio*.
 OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,

RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism
 RT distinct from that of *V. cholerae*.";
 RL Lancet 361:743-749(2003).
 DR EMBL; AP005079; BAC60103.1; -.
 KW Complete proteome.
 SQ SEQUENCE 199 AA; 23823 MW; B6275E24F8CC3D5F CRC64;

Query Match 88.2%; Score 30; DB 2; Length 199;
 Best Local Similarity 71.4%; Pred. No. 56;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
 |||:
 Db 112 FYSFATK 118

RESULT 7

Y787_HAEIN

ID Y787_HAEIN STANDARD; PRT; 201 AA.
 AC P44052;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Hypothetical protein HI0787.
 GN OrderedLocusNames=HI0787;
 OS *Haemophilus influenzae*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; *Haemophilus*.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G.G., FitzHugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 RT Rd.";
 RL Science 269:496-512(1995).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U32762; AAC22463.1; -.
 DR PIR; G64013; G64013.

DR TIGR; HI0787; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 201 AA; 23814 MW; BC4BF58FEC14DF96 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 201;
Best Local Similarity 71.4%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYXFSTK 7
|||:
Db 119 FYSFATK 125

RESULT 8

Q642R1

ID Q642R1 PRELIMINARY; PRT; 477 AA.
AC Q642R1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MGC83418 protein.
GN Name=MGC83418;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Klein S., Gerhard D.S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BC081110; AAH81110.1; -.
 SQ SEQUENCE 477 AA; 54523 MW; E42C3556EA81C4D2 CRC64;

Query Match 88.2%; Score 30; DB 2; Length 477;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
 || :|||
 Db 416 FYSYSTK 422

RESULT 9

YBFE_BACSU

ID YBFE_BACSU STANDARD; PRT; 94 AA.
 AC O31445;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Hypothetical protein ybfE.
 GN Name=ybfE; OrderedLocusNames=BSU02180;
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Haga K., Liu H., Yasumoto K., Takahashi H., Yoshikawa H.;
 RT "Sequence analysis of the 70kb region between 17 and 23 degree of the
 RT Bacillus subtilis chromosome.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
 RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
 RA Klein C., Kobayashi Y., Koetter P., Koningstein G., Krogh S.,
 RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,

RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*.";
RL Nature 390:249-256(1997).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AB006424; BAA33115.1; -.
DR EMBL; Z99105; CAB12012.1; -.
DR PIR; H69748; H69748.
DR SubtiList; BG12734; ybfE.
KW Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 5 27 Potential.
FT TRANSMEM 37 59 Potential.
SQ SEQUENCE 94 AA; 11035 MW; 7F427F191AC94B9E CRC64;

Query Match 85.3%; Score 29; DB 1; Length 94;
Best Local Similarity 71.4%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYXFSTK 7
|| |||:
Db 24 FYFFSTR 30

RESULT 10

Q65XA7

ID Q65XA7 PRELIMINARY; PRT; 100 AA.
AC Q65XA7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OJ1654_B10.11-2.
GN Name=OJ1654_B10.11-2;
OS *Oryza sativa* (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
 RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
 RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
 RA Wu H.-P., Shaw J.-F.;
 RT "Oryza sativa BAC OJ1654_B10 genomic sequence.";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AC108504; AAU44080.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 100 AA; 11996 MW; 4F72A5A7085C694A CRC64;

Query Match 85.3%; Score 29; DB 2; Length 100;
 Best Local Similarity 71.4%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
 | : | | | |
 Db 61 FFSFSTK 67

RESULT 11

Q5LXP2

ID Q6LPX2 PRELIMINARY; PRT; 202 AA.
 AC Q6LPX2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical transcriptional regulator.
 GN Name=AGR_C_43; OrderedLocusNames=PBPR2268;
 OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Photobacterium.
 OX NCBI_TaxID=74109;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
 RA Cestaro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
 RA Valle G.;
 RT "Genome analysis of Photobacterium profundum reveals the complexity of
 RT high pressure adaptations.";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: Contains 1 HTH tetR-type DNA-binding domain.
 DR EMBL; CR378670; CAG20654.1; -.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR009057; Homeodomain_like.
 DR InterPro; IPR001647; HTH_TetR.
 DR InterPro; IPR011075; TetR_like_C.
 DR Pfam; PF00440; TetR_N; 1.
 DR PRINTS; PR00455; HTH_TETR.
 KW Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.

SQ SEQUENCE 202 AA; 23081 MW; 53EB219615EADD1A CRC64;

Query Match 85.3%; Score 29; DB 2; Length 202;
Best Local Similarity 71.4%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
:| ||||
Db 45 YYFSTK 51

RESULT 12

Q63BG4

ID Q63BG4 PRELIMINARY; PRT; 207 AA.
AC Q63BG4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Transcriptional regulator, TetR family.
GN ORFNames=BTZK2162;
OS Bacillus cereus ZK.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=288681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZK;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus cereus ZK.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; CP000001; AAU18095.1; -.
SQ SEQUENCE 207 AA; 23782 MW; BE9679CCCB E0F373 CRC64;

Query Match 85.3%; Score 29; DB 2; Length 207;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
:| ||||
Db 57 YYFSTK 63

RESULT 13

Q81QL8

ID Q81QL8 PRELIMINARY; PRT; 209 AA.
AC Q81QL8; Q6HYT5; Q6KST9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Transcriptional regulator, TetR family.
GN OrderedLocusNames=BA2406, BAS2242, GBAA2406;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=Ames / isolate Porton;
 RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
 RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
 RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
 RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
 RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
 RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
 RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
 RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
 RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
 RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
 RA Salzberg S.L., Thomason B., Friedlander A.M., Koehler T.M.,
 RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
 RT "The genome sequence of *Bacillus anthracis* Ames and comparison to
 RT closely related bacteria."
 RL Nature 423:81-86(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ames / isolate 0581;
 RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
 RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
 RA Fraser C.M.;
 RT "Bacillus anthracis comparative genomics."
 RL Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sterne;
 RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: Contains 1 HTH tetR-type DNA-binding domain.
 DR EMBL; AE017031; AAP26269.1; -.
 DR EMBL; AE017334; AAT31523.1; -.
 DR EMBL; AE017225; AAT54554.1; -.
 DR TIGR; BA2406; -.
 DR TIGR; GBAA2406; -.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR009057; Homeodomain_like.
 DR InterPro; IPR001647; HTH_TetR.
 DR Pfam; PF00440; TetR_N; 1.
 DR PRINTS; PR00455; HTH_TETR.
 KW Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.
 SQ SEQUENCE 209 AA; 24026 MW; 17EE56BABF795F96 CRC64;

Query Match 85.3%; Score 29; DB 2; Length 209;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
 :| ||||
 Db 59 YYFSTK 65

Q6HIX3

ID Q6HIX3 PRELIMINARY; PRT; 209 AA.
 AC Q6HIX3;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Transcriptional regulator, TetR family.
 GN OrderedLocusNames=BT9727_2176;
 OS *Bacillus thuringiensis* (subsp. *konkukian*).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 OX NCBI_TaxID=180856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=97-27;
 RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RT "Complete genome sequence of *Bacillus thuringiensis* 97-27.";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: Contains 1 HTH tetR-type DNA-binding domain.
 DR EMBL; AE017355; AAT62138.1; -.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR009057; Homeodomain_like.
 DR InterPro; IPR001647; HTH_TetR.
 DR Pfam; PF00440; TetR_N; 1.
 DR PRINTS; PR00455; HTH_TETR.
 KW Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.
 SQ SEQUENCE 209 AA; 24052 MW; 17EE42FFAB3C5F96 CRC64;

Query Match 85.3%; Score 29; DB 2; Length 209;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
 :| |||
 Db 59 YYFSTK 65

RESULT 15

Q9YMQ2

ID Q9YMQ2 PRELIMINARY; PRT; 222 AA.
 AC Q9YMQ2;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Ld-bro-g.
 OS *Lymantria dispar* multicapsid nuclear polyhedrosis virus (LdMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10449;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99124785; PubMed=9887315; DOI=10.1006/viro.1998.9469;
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J.M., Rohrmann G.F.;

RT "Sequence and analysis of the genome of a baculovirus pathogenic for
RT Lymantria dispar.";
RL Virology 253:17-34 (1999).
DR EMBL; AF081810; AAC70261.1; -.
DR PIR; T30423; T30423.
DR InterPro; IPR003497; BRO_N.
DR Pfam; PF02498; Bro-N; 1.
SQ SEQUENCE 222 AA; 25786 MW; ECD61C41C817D49E CRC64;

Query Match 85.3%; Score 29; DB 2; Length 222;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYXFSTK 7
|||:|
Db 200 FYQFATK 206

Search completed: February 10, 2005, 15:57:33
Job time : 59.0704 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 87.1831 Seconds
(without alignments)
44.362 Million cell updates/sec

Title: US-10-067-484-5
Perfect score: 45
Sequence: 1 FYATEVXDxD 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		Length	DB	ID	Description
		Match					
1	41	91.1		10	5	ABB81972	Abb81972 30 kDa ra
2	35	77.8		13	5	ABB81974	Abb81974 30 kDa ra
3	34	75.6		346	5	ADE53197	Ade53197 FEN-1 rel
4	34	75.6		346	5	ADE53345	Ade53345 FEN-1 rel
5	34	75.6		346	7	ADA66112	Ada66112 DNAP-rela
6	34	75.6		398	6	ABU02244	Abu02244 S. pneumo
7	34	75.6		399	3	AAY81561	Aay81561 Streptoco
8	32	71.1		510	8	ADM66767	Adm66767 Listeria
9	32	71.1		865	8	ADR16230	Adr16230 Streptoco

10	32	71.1	975	8	ADS24113	Ads24113	Bacterial
11	31	68.9	92	5	ABP31657	Abp31657	Human iso
12	31	68.9	129	5	ABP07177	Abp07177	Human ORF
13	31	68.9	212	5	AAU80945	Aau80945	Haemophil
14	31	68.9	212	5	ABG94221	Abg94221	Haemophil
15	31	68.9	212	5	ABG80533	Abg80533	Haemophil
16	31	68.9	212	7	ADD24108	Add24108	Haempophi
17	31	68.9	212	7	ADJ82034	Adj82034	Protein f
18	31	68.9	212	7	ADK17122	Adk17122	Virus-lik
19	31	68.9	229	4	AAU67906	Aau67906	Propionib
20	31	68.9	229	6	ABM64425	Abm64425	Propionib
21	31	68.9	243	7	ABO61951	Abo61951	Klebsiell
22	31	68.9	381	6	ABU22905	Abu22905	Protein e
23	31	68.9	464	6	AAE38286	Aae38286	Rice enha
24	31	68.9	516	6	ABU25295	Abu25295	Protein e
25	31	68.9	823	4	AAU41924	Aau41924	Propionib
26	31	68.9	823	6	ABM38443	Abm38443	Propionib
27	31	68.9	961	6	ABM66081	Abm66081	Propionib
28	31	68.9	971	4	AAU50418	Aau50418	Propionib
29	31	68.9	971	6	ABM46937	Abm46937	Propionib
30	31	68.9	1040	6	ABM65765	Abm65765	Propionib
31	30	66.7	44	8	ADJ56936	Adj56936	HIV-1 env
32	30	66.7	44	8	ADR58152	Adr58152	Novel ant
33	30	66.7	123	8	ADP29814	Adp29814	Human sec
34	30	66.7	142	5	ABP00377	Abp00377	Human ORF
35	30	66.7	229	4	AAU36798	Aau36798	Staphyloc
36	30	66.7	229	6	ABU16090	Abu16090	Protein e
37	30	66.7	229	6	ABM71827	Abm71827	Staphyloc
38	30	66.7	277	7	ADC87255	Adc87255	Human GPC
39	30	66.7	408	6	ABU25387	Abu25387	Protein e
40	30	66.7	422	3	AAU91060	Aay91060	Streptomy
41	30	66.7	433	2	AAW02649	Aaw02649	Ascorbate
42	30	66.7	439	7	ABO61981	Abo61981	Klebsiell
43	30	66.7	441	3	AAG38523	Aag38523	Arabidops
44	30	66.7	446	6	ABO27178	Abo27178	Human sig
45	30	66.7	508	6	ABU24797	Abu24797	Protein e

ALIGNMENTS

RESULT 1

ABB81972

ID ABB81972 standard; peptide; 10 AA.

XX

AC ABB81972;

XX

DT 25-NOV-2002 (first entry)

XX

DE 30 kDa ragweed pollen allergen tryptic peptide 5.

XX

KW Ragweed; pollen; allergen; Ambt 7; glycoprotein; antiallergic;
KW immunotherapy; disulphide protein.

XX

OS Ambrosia elatior.

XX

FH Key Location/Qualifiers

FT Misc-difference 7
 FT /label= Leu or Ile
 FT Misc-difference 9
 FT /label= Leu or Ile
 XX
 PN WO200263012-A2.
 XX
 PD 15-AUG-2002.
 XX
 PF 04-FEB-2002; 2002WO-US003346.
 XX
 PR 05-FEB-2001; 2001US-0266686P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Buchanan BB, Del Val G, Frick OL;
 XX
 DR WPI; 2002-657539/70.
 XX
 PT New ragweed pollen allergens, useful in allergy testing and immunotherapy
 PT regimens, particularly for treating sensitivity to pollen or pollen
 PT allergy (e.g. anaphylaxis, or symptoms of hives or asthma) in a mammal,
 PT especially a human.
 XX
 PS Claim 1; Page 53; 70pp; English.
 XX
 CC The invention relates to an isolated pollen allergen purified from
 CC ragweed pollen, substantially free of any other pollen proteins, or a
 CC protein that is an antigenic fragment of a pollen allergen Ambt 7. The
 CC allergen is characterized by the following physiochemical and biological
 CC properties: (a) being contained in pollen extracts; (b) a glycoprotein;
 CC (c) a sulphhydryl group containing protein; (d) a molecular weight of
 CC about 30 kDa as determined by SDS-polyacrylamide gel electrophoresis; and
 CC (e) possessing allergen activity. The pollen allergen, or antigenic
 CC protein fragment of the pollen allergen Ambt 7, or composition is useful
 CC for treating sensitivity to pollen or pollen allergy in a mammal. This
 CC allergy includes anaphylaxis or atopy, which includes the symptoms of hay
 CC fever, asthma or hives. The allergen is also useful in allergy testing
 CC and immunotherapy regimens. Sequences ABB81968-978 represent tryptic
 CC peptide fragments of the 30 kDa ragweed complete pollen extract
 CC disulphide protein allergen
 XX
 SQ Sequence 10 AA;

Query Match 91.1%; Score 41; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.06;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
 |||||
 Db 1 FYATEVXDxD 10

RESULT 2
 ABB81974
 ID ABB81974 standard; peptide; 13 AA.
 XX

AC ABB81974;
 XX
 DT 25-NOV-2002 (first entry)
 XX
 DE 30 kDa ragweed pollen allergen tryptic peptide 7.
 XX
 KW Ragweed; pollen; allergen; Ambt 7; glycoprotein; antiallergic;
 KW immunotherapy; disulphide protein.
 XX
 OS Ambrosia elatior.
 XX
 PN WO200263012-A2.
 XX
 PD 15-AUG-2002.
 XX
 PF 04-FEB-2002; 2002WO-US003346.
 XX
 PR 05-FEB-2001; 2001US-0266686P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Buchanan BB, Del Val G, Frick OL;
 XX
 DR WPI; 2002-657539/70.
 XX
 PT New ragweed pollen allergens, useful in allergy testing and immunotherapy
 PT regimens, particularly for treating sensitivity to pollen or pollen
 PT allergy (e.g. anaphylaxis, or symptoms of hives or asthma) in a mammal,
 PT especially a human.
 XX
 PS Claim 1; Page 53; 70pp; English.
 XX
 CC The invention relates to an isolated pollen allergen purified from
 CC ragweed pollen, substantially free of any other pollen proteins, or a
 CC protein that is an antigenic fragment of a pollen allergen Ambt 7. The
 CC allergen is characterized by the following physiochemical and biological
 CC properties: (a) being contained in pollen extracts; (b) a glycoprotein;
 CC (c) a sulphydryl group containing protein; (d) a molecular weight of
 CC about 30 kDa as determined by SDS-polyacrylamide gel electrophoresis; and
 CC (e) possessing allergen activity. The pollen allergen, or antigenic
 CC protein fragment of the pollen allergen Ambt 7, or composition is useful
 CC for treating sensitivity to pollen or pollen allergy in a mammal. This
 CC allergy includes anaphylaxis or atopy, which includes the symptoms of hay
 CC fever, asthma or hives. The allergen is also useful in allergy testing
 CC and immunotherapy regimens. Sequences ABB81968-978 represent tryptic
 CC peptide fragments of the 30 kDa ragweed complete pollen extract
 CC disulphide protein allergen
 XX
 SQ Sequence 13 AA;

Query Match 77.8%; Score 35; DB 5; Length 13;
 Best Local Similarity 77.8%; Pred. No. 1.5;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YATEVXDxD 10
 |||||
 Db 2 YATEVLdLD 10

RESULT 3

ADE53197

ID ADE53197 standard; protein; 346 AA.

XX

AC ADE53197;

XX

DT 29-JAN-2004 (first entry)

XX

DE FEN-1 related polypeptide used within the scope of the invention, #33.

XX

KW Flap endonuclease-1; FEN-1; endonuclease; structure-specific nuclease;
KW invasive cleavage structure; thermostable; DNA polymerase; 5' nuclease;
KW viral infection; bacterial infection; cancer; forensic analysis;
KW paternity determination.

XX

OS *Pyrobaculum aerophilum*.

XX

PN WO200270755-A2.

XX

PD 12-SEP-2002.

XX

PF 15-NOV-2001; 2001WO-US044953.

XX

PR 15-NOV-2000; 2000US-00713601.

PR 17-NOV-2000; 2000US-00714935.

XX

PA (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX

PI Lyamichev VI, Kaiser MW, Lyamicheva N;

XX

DR WPI; 2002-750464/81.

DR N-PSDB; ADE53196.

XX

PT New composition useful for detecting and characterizing nucleic acid
PT sequences and sequence variants for detecting the presence of viral or
PT bacterial infections or cancer, comprises purified or chimerical FEN-1
PT endonuclease.

XX

PS Claim 12; SEQ ID NO 379; 871pp; English.

XX

CC The invention discloses a new composition (I) which comprises a purified
CC flap endonuclease-1 (FEN-1) from e.g. *Sulfolobus solfataricus*,
CC *Pyrobaculum aerophilum* or a chimerical FEN-1 endonuclease having a
CC portion of the above endonuclease in addition to that of *Pyrococcus*
CC *horikoshii* and *Aeropyrum pernix*. Also claimed is a composition comprising
CC an isolated nucleic acid sequence encoding the endonuclease mentioned
CC above, a composition comprising a vector having the nucleic acid sequence
CC cited above, a composition comprising a host cell and vector cited above,
CC a mixture comprising a first structure-specific nuclease selected from
CC the species mentioned in composition (I), and a purified second structure
CC -specific nuclease and detecting a target sequence, comprising: (a)
CC providing a sample suspected of containing the target sequence,
CC oligonucleotides capable of forming an invasive cleavage structure in the
CC presence of the target sequence, and a FEN-1 endonuclease selected from
CC the species cited above and (b) exposing the sample to the

CC oligonucleotides and FEN-1 endonuclease. The second structure-specific
 CC nuclease also comprises a thermostable DNA polymerase. It has a 5'
 CC nuclease derived from a DNA polymerase altered in amino acid sequence
 CC such that it exhibits reduced DNA synthetic activity from that of the
 CC wild-type DNA polymerase but retains substantially the same 5' nuclease
 CC activity of the wild-type DNA polymerase. The second structure is
 CC selected from CLEAVASE BN enzyme, CLEAVASE DA enzyme, CLEAVASE DN enzyme,
 CC CLEAVASE DV enzyme, CLEAVASE BN/thrombin enzyme, CLEAVASE TThDN enzyme,
 CC T. aquaticus DNA polymerase, T. thermophilus DNA polymerase, E. coli Exo
 CC III and S. cerevisiae Rad1/Rad10 complex. The nucleic acid treatment kit
 CC comprises (I) and oligonucleotides capable of forming an invasive
 CC cleavage structure in the presence of a target nucleic acid. The
 CC oligonucleotides comprise: (a) a first oligonucleotide having a 5'
 CC portion complementary to a first portion of a target nucleic acid and (b)
 CC a second oligonucleotide comprising a 5' portion complementary to a
 CC second portion of the target nucleic acid downstream of and contiguous to
 CC the first portion and a 3' portion. The 3' portion of the second
 CC oligonucleotide comprises a single 3' terminal nucleotide not
 CC complementary to the target nucleic acid. Additionally, the kit has a
 CC third oligonucleotide complementary to a third portion of the target
 CC nucleic acid upstream of the first portion of the first target nucleic
 CC acid. In detecting a target sequence, the oligonucleotides and
 CC endonuclease are mixed under conditions where an invasive cleavage
 CC structure is formed between the target sequence and the oligonucleotides
 CC if the target sequence is present in the sample, where the invasive
 CC cleavage structure is cleaved by the endonuclease to form a cleavage
 CC product. The composition is useful in detecting and characterising
 CC specific nucleic acid sequences and sequence variants which can be used
 CC in detecting the presence of viral or bacterial infections, and other
 CC diseases such as cancer. The composition may also be used in forensic
 CC analysis or for paternity determinations. The sequence presented is a FEN
 CC -1 related polypeptide used within the scope of the invention.

XX

SQ Sequence 346 AA;

Query Match 75.6%; Score 34; DB 5; Length 346;
 Best Local Similarity 77.8%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
 |||||
 Db 296 YATEVRDPD 304

RESULT 4

ADE53345

ID ADE53345 standard; protein; 346 AA.

XX

AC ADE53345;

XX

DT 29-JAN-2004 (first entry)

XX

DE FEN-1 related polypeptide used within the scope of the invention, #55.

XX

KW Flap endonuclease-1; FEN-1; endonuclease; structure-specific nuclease;
 KW invasive cleavage structure; thermostable; DNA polymerase; 5' nuclease;
 KW viral infection; bacterial infection; cancer; forensic analysis;

KW paternity determination.
 XX
 OS *Pyrobaculum aerophilum*.
 XX
 PN WO200270755-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 15-NOV-2001; 2001WO-US044953.
 XX
 PR 15-NOV-2000; 2000US-00713601.
 PR 17-NOV-2000; 2000US-00714935.
 XX
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
 XX
 PI Lyamichev VI, Kaiser MW, Lyamicheva N;
 XX
 DR WPI; 2002-750464/81.
 XX
 PT New composition useful for detecting and characterizing nucleic acid
 PT sequences and sequence variants for detecting the presence of viral or
 PT bacterial infections or cancer, comprises purified or chimerical FEN-1
 PT endonuclease.
 XX
 PS Disclosure; SEQ ID NO 527; 871pp; English.
 XX
 CC The invention discloses a new composition (I) which comprises a purified
 CC flap endonuclease-1 (FEN-1) from e.g. *Sulfolobus solfataricus*,
 CC *Pyrobaculum aerophilum* or a chimerical FEN-1 endonuclease having a
 CC portion of the above endonuclease in addition to that of *Pyrococcus*
 CC *horikoshii* and *Aeropyrum pernix*. Also claimed is a composition comprising
 CC an isolated nucleic acid sequence encoding the endonuclease mentioned
 CC above, a composition comprising a vector having the nucleic acid sequence
 CC cited above, a composition comprising a host cell and vector cited above,
 CC a mixture comprising a first structure-specific nuclease selected from
 CC the species mentioned in composition (I), and a purified second structure
 CC -specific nuclease and detecting a target sequence, comprising: (a)
 CC providing a sample suspected of containing the target sequence,
 CC oligonucleotides capable of forming an invasive cleavage structure in the
 CC presence of the target sequence; and a FEN-1 endonuclease selected from
 CC the species cited above and (b) exposing the sample to the
 CC oligonucleotides and FEN-1 endonuclease. The second structure-specific
 CC nuclease also comprises a thermostable DNA polymerase. It has a 5'
 CC nuclease derived from a DNA polymerase altered in amino acid sequence
 CC such that it exhibits reduced DNA synthetic activity from that of the
 CC wild-type DNA polymerase but retains substantially the same 5' nuclease
 CC activity of the wild-type DNA polymerase. The second structure is
 CC selected from CLEAVASE BN enzyme, CLEAVASE DA enzyme, CLEAVASE DN enzyme,
 CC CLEAVASE DV enzyme, CLEAVASE BN/thrombin enzyme, CLEAVASE TThDN enzyme,
 CC *T. aquaticus* DNA polymerase, *T. thermophilus* DNA polymerase, *E. coli* Exo
 CC III and *S. cerevisiae* Rad1/Rad10 complex. The nucleic acid treatment kit
 CC comprises (I) and oligonucleotides capable of forming an invasive
 CC cleavage structure in the presence of a target nucleic acid. The
 CC oligonucleotides comprise: (a) a first oligonucleotide having a 5'
 CC portion complementary to a first portion of a target nucleic acid and (b)
 CC a second oligonucleotide comprising a 5' portion complementary to a
 CC second portion of the target nucleic acid downstream of and contiguous to

CC the first portion and a 3' portion. The 3' portion of the second
 CC oligonucleotide comprises a single 3' terminal nucleotide not
 CC complementary to the target nucleic acid. Additionally, the kit has a
 CC third oligonucleotide complementary to a third portion of the target
 CC nucleic acid upstream of the first portion of the first target nucleic
 CC acid. In detecting a target sequence, the oligonucleotides and
 CC endonuclease are mixed under conditions where an invasive cleavage
 CC structure is formed between the target sequence and the oligonucleotides
 CC if the target sequence is present in the sample, where the invasive
 CC cleavage structure is cleaved by the endonuclease to form a cleavage
 CC product. The composition is useful in detecting and characterising
 CC specific nucleic acid sequences and sequence variants which can be used
 CC in detecting the presence of viral or bacterial infections, and other
 CC diseases such as cancer. The composition may also be used in forensic
 CC analysis or for paternity determinations. The sequence presented is a FEN
 CC -1 related polypeptide used within the scope of the invention.
 XX
 SQ Sequence 346 AA;

Query Match 75.6%; Score 34; DB 5; Length 346;
 Best Local Similarity 77.8%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
 ||||| |
 Db 296 YATEVRDPD 304

RESULT 5

ADA66112

ID ADA66112 standard; protein; 346 AA.
 XX
 AC ADA66112;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE DNAP-related protein #1.
 XX
 KW DNAP; invasive cleavage structure; dendrimer; nuclease; endonuclease;
 KW polymerase; polyglycol; 5' nuclease; allelic variation.
 XX
 OS Unidentified;.
 XX
 PN US2003044796-A1.
 XX
 PD 06-MAR-2003.
 XX
 PF 27-AUG-2001; 2001US-00940244.
 XX
 PR 26-NOV-1996; 96US-00756386.
 PR 24-MAR-1998; 98WO-US005809.
 PR 09-JUL-1999; 99US-00350309.
 PR 08-FEB-2000; 2000US-00381212.
 PR 08-DEC-2000; 2000US-00732622.
 XX
 PA (NERI/) NERI B P.
 PA (HALL/) HALL J G.

PA (LYAM/) LYAMICHEV V.
 PA (SMIT/) SMITH L M.
 XX
 PI Neri BP, Hall JG, Lyamichev V, Smith LM;
 XX
 DR WPI; 2003-596420/56.
 DR N-PSDB; ADA66261.
 XX
 PT Detection system for nucleic acid sequences comprises oligonucleotides
 PT configured for hybridizing to target nucleic acid to form invasive
 PT cleavage structure and dendrimer.
 XX
 PS Disclosure; Fig 155; 354pp; English.
 XX
 CC The invention relates to a detection system which has oligonucleotides
 CC configured for hybridisation to a target nucleic acid to form an invasive
 CC cleavage structure and dendrimer, where the oligonucleotides are attached
 CC to the dendrimer. The invention also relates to a method for
 CC characterising a nucleic acid sequence comprising providing a sample
 CC suspected of containing a target nucleic acid, oligonucleotides
 CC configured to hybridise to the target nucleic acid to form an invasive
 CC cleavage structure and a dendrimer to which the oligonucleotide is
 CC attached, and exposing the sample to the oligonucleotides and an agent
 CC that detects the presence of an invasive cleavage structure. The agent
 CC comprises a cleavage agent having a structure-specific nuclease,
 CC preferably a 5' nuclease comprising an endonuclease or polymerase. The
 CC detection system further comprises a spacer molecule, consisting of a
 CC carbon chain, polynucleotide or polyglycol, to which the oligonucleotide
 CC is attached. The target molecule and the agent are attached to a solid
 CC support. The invention is used in the detection and characterisation of
 CC nucleic acid sequences and variations in these sequences, used in allelic
 CC variation studies. This sequence represents a protein used in the scope
 CC of the invention.
 XX
 SQ Sequence 346 AA;

Query Match 75.6%; Score 34; DB 7; Length 346;
 Best Local Similarity 77.8%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YATEVXDxD 10
 |||||
 Db 296 YATEVRDPD 304

RESULT 6

ABU02244

ID ABU02244 standard; protein; 398 AA.

XX

AC ABU02244;

XX

DT 23-OCT-2003 (revised)

DT 11-FEB-2003 (first entry)

XX

DE S. pneumoniae type 4 strain protein from coding region #1822.

XX

KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;

KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX
 OS Streptococcus pneumoniae; type 4 strain.
 XX
 PN WO200277021-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 27-MAR-2002; 2002WO-IB002163.
 XX
 PR 27-MAR-2001; 2001GB-00007658.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Massignani V, Tettelin H, Fraser C;
 XX
 DR WPI; 2003-040579/03.
 DR N-PSDB; ABX07534.
 XX
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 PT ear infection.
 XX
 PS Claim 1; SEQ ID NO 3644; 56pp; English.
 XX
 CC The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS56454. Also included are an antibody which binds one of the proteins,
 CC treating a patient by administering the protein, DNA or antibody (in a
 CC composition), a kit comprising first and second primers, which are the
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a
 CC sequence not defined in the specification, for amplifying a target
 CC sequence contained within a Streptococcus nucleic acid sequence, where
 CC the first primer is substantially complementary to the target sequence
 CC and the second primer is substantially complementary to the complement of
 CC the target sequence, and where the parts of the primers having
 CC substantial complementarity define the termini of the target sequence to
 CC be amplified, assay comprising contacting a test compound with the
 CC protein, and determining whether the test compound binds to the protein
 CC and a Streptococcus pneumoniae bacterium, where one or more genes
 CC encoding the proteins has been rendered inactive. The proteins, nucleic
 CC acid molecules, antibody and compositions are useful as medicaments for
 CC treating or preventing a disease or infection due to streptococcus
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
 CC media or ear infection. They are also useful in developing vaccines,
 CC diagnostics and antibiotics. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to

CC standardise OS field)
XX
SQ Sequence 398 AA;

Query Match 75.6%; Score 34; DB 6; Length 398;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
|:|||| : |
Db 86 FFATEVVESD 95

RESULT 7

AAy81561

ID AAY81561 standard; protein; 399 AA.

XX

AC AAY81561;

XX

DT 24-MAY-2000 (first entry)

XX

DE Streptococcus pneumoniae type 4 protein sequence #61.

XX

KW Streptococcus pneumoniae; vaccine; screening; protein antigen;
KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
KW pneumococcal disease.

XX

OS Streptococcus pneumoniae.

XX

PN WO200006737-A2.

XX

PD 10-FEB-2000.

XX

PF 27-JUL-1999; 99WO-GB002451.

XX

PR 27-JUL-1998; 98GB-00016337.

PR 19-MAR-1999; 99US-0125164P.

XX

PA (MICR-) MICROBIAL TECHNIQS LTD.

XX

PI Gilbert CFG, Hansbro PM;

XX

DR WPI; 2000-195300/17.

XX

PT New Streptococcal protein, useful as a vaccine, for diagnosis of
PT pneumococcal diseases and for screening agents capable of antagonizing or
PT inhibiting expression of the protein.

XX

PS Claim 1; Page 78; 108pp; English.

XX

CC AAY81501 to AAY81679 represent specifically claimed protein sequences
CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent
CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
CC The sequences have antibacterial and antiinflammatory properties. The
CC protein sequences, and fragments of them, are useful as immunogens and/or
CC antigens. The nucleotide sequences can be used in vaccines and in
CC diagnostic assays. The proteins and nucleotides can be useful for the

CC detection and diagnosis of *S. pneumoniae*. The protein sequences are also
CC useful for screening an agent capable of antagonising, inhibiting or
CC interfering with the function or expression of the proteins in which the
CC agent is useful for treatment or prophylaxis of *S. pneumoniae* infection
CC and meningitis. AAA05591 to AAA05614 represent primers used in the
CC exemplification of the present invention

XX

SQ Sequence 399 AA;

Query Match 75.6%; Score 34; DB 3; Length 399;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYATEVXDXD 10
|:||||:|
Db 86 FFATEVVESD 95

RESULT 8

ADM66767

ID ADM66767 standard; protein; 510 AA.

XX

AC ADM66767;

XX

DT 03-JUN-2004 (first entry)

XX

DE *Listeria* thermolysin-like protease (TLP) precursor protein 1.

XX

KW thermolysin-like protease; TLP; S1' site; gluten degradation; wheat;

KW baking industry; beer clarification; brewing; dehairing; skin dewooling;

KW leather; protein hydrolysate production; artificial sweetener; aspartame;

KW precursor; enzyme.

XX

OS *Listeria*.

XX

PN WO2004011619-A2.

XX

PD 05-FEB-2004.

XX

PF 28-JUL-2003; 2003WO-US023726.

XX

PR 26-JUL-2002; 2002US-0398656P.

XX

PA (STRA-) STRATAGENE.

XX

PI Clark DD, Braman JC;

XX

DR WPI; 2004-143847/14.

XX

PT New thermolysin-like protease with substrate specificity for a basic or

PT an acidic amino acid, useful in biological and biomedical research,

PT identifying therapeutic agents and diagnostic markers, or producing

PT artificial sweeteners.

XX

PS Disclosure; SEQ ID NO 60; 82pp; English.

XX

CC The invention relates to a novel thermolysin-like protease (TLP)

CC comprising an S1' site and modified to have a substrate specificity for a
CC basic or an acidic amino acid. The thermolysin-like protease of the
CC invention may be useful in proteolysis applications, biological and
CC biomedical research, identifying therapeutic agents and diagnostic
CC markers, characterising cells and organisms that have undergone genetic
CC modifications, identifying unknown illnesses, characterising polypeptides
CC or identifying biological samples. The thermolysin-like protease may also
CC be useful in industrial processes, such as the degradation of gluten from
CC wheat within the baking industry, clarification of beer within the
CC brewing industry, dehairing or dewooling of skins within the leather
CC industry, preparation of protein hydrolysates or production of artificial
CC sweeteners like aspartame. The current sequence is that of a TLP
CC precursor protein of the invention.

XX

SQ Sequence 510 AA;

Query Match 71.1%; Score 32; DB 8; Length 510;
Best Local Similarity 75.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8
|||:| |
Db 282 FYASEVYD 289

RESULT 9

ADR16230

ID ADR16230 standard; protein; 865 AA.

XX

AC ADR16230;

XX

DT 21-OCT-2004 (first entry)

XX

DE Streptococcus pyogenes serum opacity factor 61 (sof61) partial protein.

XX

KW Serum opacity factor; SOF; fibronectin binding protein; FnBA;
KW streptococcal infection; acute rheumatic fever; acute glomerulonephritis;
KW associated autoimmune neurological disorder; antibacterial; therapy.

XX

OS Streptococcus pyogenes.

XX

PN US2004151737-A1.

XX

PD 05-AUG-2004.

XX

PF 04-FEB-2004; 2004US-00771931.

XX

PR 05-FEB-2003; 2003US-0446061P.

XX

PA (UYTE-) UNIV TENNESSEE.

XX

PI Courtney HS;

XX

DR WPI; 2004-561475/54.

DR GENBANK; AF138804.

XX

PT New composition comprising immunogenic portions from Group A streptococci

PT serum opacity factors (SOF) or Group C streptococci fibronectin binding
PT protein (FnBA), useful for treating, preventing, or monitoring of
PT streptococcal infections.
XX
PS Claim 9; SEQ ID NO 43; 89pp; English.
XX
CC The present invention provides Group A streptococci serum opacity factors
CC (SOF) proteins, Group C streptococci fibronectin binding protein (FnBA),
CC their encoding polynucleotides and their immunogenic epitopes. The
CC invention is useful for eliciting an opsonic and/or protective antibodies
CC specific for Streptococcus pyogenes and/or Streptococcus dysgalactiae.
CC The invention is also useful for treating, preventing and monitoring
CC streptococcal infections such as acute rheumatic fever, acute
CC glomerulonephritis and associated autoimmune neurological disorders. The
CC invention acts as an antibacterial agent. The present sequence is
CC Streptococcus pyogenes serum opacity factor (SOF) partial protein.
XX
SQ Sequence 865 AA;

Query Match 71.1%; Score 32; DB 8; Length 865;
Best Local Similarity 50.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYATEVXDxD 10
||: :| | |
Db 544 FYSVDVTDS 553

RESULT 10

ADS24113

ID ADS24113 standard; protein; 975 AA.

XX

AC ADS24113;

XX

DT 02-DEC-2004 (first entry)

XX

DE Bacterial polypeptide #13146.

XX

KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX

OS Bacteria.

XX

PN US2003233675-A1.

XX

PD 18-DEC-2003.

XX

PF 20-FEB-2003; 2003US-00369493.

XX

PR 21-FEB-2002; 2002US-0360039P.

XX

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 13146; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 975 AA;

Query Match 71.1%; Score 32; DB 8; Length 975;
 Best Local Similarity 66.7%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
 | | | | |
 Db 933 YVTEVSDLD 941

RESULT 11
 ABP31657
 ID ABP31657 standard; protein; 92 AA.
 XX
 AC ABP31657;
 XX

DT 08-JUL-2002 (first entry)

XX

DE Human isomerase-like ORF630 protein, SEQ ID NO:1260.

XX

KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;

KW disease monitoring; cytokine; cell proliferation; cell differentiation;

KW immune modulation; haematopoiesis regulation; tissue growth;

KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;

KW thrombolytic; tumour inhibition; bodily characteristic; fertility;

KW behaviour; cancer; proliferative disorder; neurological disorder;

KW cardiovascular disease; immune system disorder; organ transplantation;

KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;

KW hypothyroidism; cholesterol ester storage disease; infection; vulnerary;

KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;

KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;

KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;

KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX

OS Homo sapiens.

XX

PN WO200190366-A2.

XX

PD 29-NOV-2001.

XX

PF 24-MAY-2001; 2001WO-US017076.

XX

PR 24-MAY-2000; 2000US-0206690P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Leach MD, Shimkets RA;

XX

DR WPI; 2002-106200/14.

DR N-PSDB; ABN75683.

XX

PT Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,

PT hyperproliferative disorders and disorders related to organ

PT transplantation.

XX

PS Claim 10; Page 580; 2508pp; English.

XX

CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins

CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-

CC ABN79587 represent cDNAs encoding them. The invention also encompasses

CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively

CC referred to as ORFX) proteins, polynucleotides at least 85% identical to

CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX

CC polynucleotides, the recombinant production of ORFX proteins, antibodies

CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and

CC polypeptides, methods of screening for modulators of ORFX expression or

CC activity, and methods of screening individuals for a predisposition to an

CC ORFX-associated disorder. The ORFX proteins of the invention have a wide

CC range of biological activities, such as cytokine, cell proliferation,

CC cell differentiation, immune modulation, haematopoiesis regulation,

CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/

CC chemokinetic activity, haemostatic activity, thrombolytic activity,

CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein; and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases

XX

SQ Sequence 92 AA;

Query Match 68.9%; Score 31; DB 5; Length 92;
 Best Local Similarity 66.7%; Pred. No. 93;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YATEVXDxD 10
 | | | | |
 Db 56 YVTEVLDDD 64

RESULT 12

ABP07177

ID ABP07177 standard; protein; 129 AA.

XX

AC ABP07177;

XX

DT 25-JUN-2002 (first entry)

XX

DE Human ORFX protein sequence SEQ ID NO:14336.

XX

KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

XX

OS Homo sapiens.

XX

PN WO200192523-A2.

XX

PD 06-DEC-2001.

XX

PF 29-MAY-2001; 2001WO-US010836.

XX
PR 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI; 2002-106308/14.
DR N-PSDB; ABN22929.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
PS Disclosure; SEQ ID NO 14336; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table.1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 129 AA;

Query Match 68.9%; Score 31; DB 5; Length 129;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 FYATEVXDxD 10
||| | | |
Db 78 FYANTVTDLD 87

RESULT 13
AAU80945
ID AAU80945 standard; protein; 212 AA.
XX

AC AAU80945;
XX
DT 09-APR-2002 (first entry)
XX
DE Haemophilus influenzae pilin protein.
XX
KW Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW human immunodeficiency virus infection; viral hepatitis; measles;
KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy; cancer;
KW chronic disease; arthritis; colitis; diabetes; multiple sclerosis.
XX
OS Haemophilus influenzae.
XX
PN WO200185208-A2.
XX
PD 15-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-IB000741.
XX
PR 05-MAY-2000; 2000US-0202341P.
XX
PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA (SEBB/) SEBBEL P.
PA (DUNA/) DUNANT N.
PA (BACH/) BACHMANN M.
PA (TISS/) TISSOT A.
PA (LECH/) LECHENER F.
XX
PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechener F;
XX
DR WPI; 2002-055561/07.
XX
PT New composition, useful for vaccine production, comprises antigen or
PT antigenic determinant and non-natural molecular scaffold comprising
PT organizer and core particle such as bacterial pilus or pilin protein.
XX
PS Disclosure; Page 248; 287pp; English.
XX
CC The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold (molecular scaffold) which comprises a core particle
CC such as a bacterial pilus or pilin protein, a recombinant form of the
CC protein, a virus-like particle or a hepatitis B virus capsid protein
CC (HBcAg), and an organiser; and (b) an antigen or antigenic determinant,
CC where the molecular scaffold and antigenic determinant interact to form
CC an ordered and repetitive antigen array. Suitable antigenic determinants
CC include JUN, FOS, HIV gp140, measles virus N protein, bee venom
CC phospholipase, Sinbis virus E2 protein, amyloid beta derived peptides and
CC influenza M2 antigen. The composition (or vaccine) is useful for
CC immunisation, by administration to a subject, where the administration
CC produces an immune response, such as humoral, cellular or protective
CC immune response, preferably a Th type 2 T-helper (Th2) response that is
CC specific for the antigenic determinant. The administration induces
CC antibodies specific for the antigenic determinant of a subtype
CC corresponding to the Th2 subtype in the subject. The subject does not

CC generate a Th2 subtype that is specific for pilus or pilin polypeptide or
CC antigenic determinant. The composition is useful for the production of
CC vaccines for prevention of infectious diseases such as human
CC immunodeficiency virus, viral hepatitis, measles, chicken pox, pneumonia,
CC tuberculosis, syphilis, malaria, and for treating allergy, cancer, and
CC chronic diseases induced or accelerated by a Th1 type immune response,
CC such as arthritis, colitis, diabetes and multiple sclerosis. The
CC composition is useful to generate defined self-specific antibodies and
CC specific immune responses of the Th2 type and allows the creation of
CC highly efficient vaccines against infectious diseases, and for treating
CC allergy, cancer, and chronic diseases induced or accelerated by a Th1
CC type immune response. The present sequence is a peptide or protein
CC incorporated into the compositions of the invention

XX

SQ Sequence 212 AA;

Query Match 68.9%; Score 31; DB 5; Length 212;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
||: |: ||
Db 98 FYSWEIADKD 107

RESULT 14

ABG94221

ID ABG94221 standard; protein; 212 AA.

XX

AC ABG94221;

XX

DT 10-DEC-2002 (first entry)

XX

DE Haemophilus influenzae pillin protein.

XX

KW Human; mouse; rat; antimicrobial; antiallergic; immunomodulatory;
KW cytostatic; antiviral; antidiabetic; hypoglycaemic; antigen array;
KW vaccine; infectious disease.

XX

OS Haemophilus influenzae.

XX

PN WO200256905-A2.

XX

PD 25-JUL-2002.

XX

PF 21-JAN-2002; 2002WO-IB000166.

XX

PR 19-JAN-2001; 2001US-0262379P.

PR 04-MAY-2001; 2001US-0288549P.

PR 05-OCT-2001; 2001US-0326998P.

PR 07-NOV-2001; 2001US-0331045P.

XX

PA (CYTO-) CYTOS BIOTECHNOLOGY AG.

XX

PI Renner WA, Bachmann M, Tissot A, Maurer P, Lechner F, Sebbel P;
PI Piossek C;

XX

DR WPI; 2002-627351/67.
 XX
 PT Molecular antigen array used in the production of vaccines for infectious
 PT diseases.
 XX
 PS Disclosure; Page 369-370; 441pp; English.
 XX
 CC This invention relates to a novel ordered and repetitive antigen array
 CC used in the production of vaccines for infectious diseases. The invention
 CC also discloses a composition comprising a non-natural molecular scaffold
 CC comprising a core particle selected from a core particle of a non-natural
 CC origin and a core particle of natural origin and an organiser comprising
 CC at least one first attachment site, where the organiser is connected to
 CC the core particle by at least one covalent bond. Also disclosed is an
 CC antigen or antigenic determinant with at least one second attachment
 CC site, where the antigen or antigenic determinant is amyloid beta peptide
 CC (Abeta1-42) or its fragment and where the second attachment site is
 CC selected from an attachment site not naturally occurring with the antigen
 CC or antigenic determinant and an attachment site naturally occurring with
 CC the antigen or antigenic determinant, where the second attachment site is
 CC capable of association through at least one non-peptide bond to the first
 CC attachment site and where the antigen or antigenic determinant and the
 CC scaffold interact through the association to form an ordered and
 CC repetitive antigen array. The invention also comprises a coat protein
 CC capable of forming a capsid which comprises mutant Qbeta coat proteins
 CC having an amino acid sequence selected from five amino acid sequences
 CC fully defined in the specification. The compounds of the invention may
 CC have antimicrobial, antiallergic, immunomodulatory, cytostatic,
 CC antiviral, antidiabetic, or hypoglycaemic activities and may be used in
 CC immunisation and as a vaccine. The present sequence represents a protein
 CC sequence used to create the compositions of the invention
 XX
 SQ Sequence 212 AA;

Query Match 68.9%; Score 31; DB 5; Length 212;
 Best Local Similarity 50.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYATEVXDxD 10
 ||: |: ||
 Db 98 FYSWEIADKD 107

RESULT 15

ABG80533

ID ABG80533 standard; protein; 212 AA.
 XX
 AC ABG80533;
 XX
 DT 29-NOV-2002 (first entry)
 XX
 DE Haemophilus influenzae pilin protein.
 XX
 KW Molecular antigen array; vaccine; antigen; antimicrobial;
 KW molecular scaffold; amyloid beta; Abeta 1-42; influenza;
 KW graft versus host disease; IgE-mediated allergic reaction; anaphylaxis;
 KW adult respiratory distress syndrome; ARDS; Crohn's disease;

KW allergic asthma; acute lymphoblastic leukaemia; non-Hodgkin's lymphoma;
KW Grave's disease; systemic lupus erythematosus; osteoporosis;
KW inflammatory immune disease; myasthenia gravis; multiple sclerosis;
KW immunoproliferative disease lymphadenopathy; Alzheimer's disease;
KW angioimmunoproliferative lymphadenopathy; immunoblastic lymphadenopathy;
KW rheumatoid arthritis; diabetes; infectious disease; factor Xa;
KW enterokinase; cysteine-containing linker.

XX

OS Haemophilus influenzae.

XX

PN WO200256907-A2.

XX

PD 25-JUL-2002.

XX

PF 21-JAN-2002; 2002WO-IB000168.

XX

PR 19-JAN-2001; 2001US-0262379P.

PR 04-MAY-2001; 2001US-0288549P.

PR 05-OCT-2001; 2001US-0326998P.

PR 07-NOV-2001; 2001US-0331045P.

XX

PA (CYTO-) CYTOS BIOTECHNOLOGY AG.

PA (NOVS) NOVARTIS PHARMA AG.

PA (MAUR/) MAURER P.

PA (LECH/) LECHNER F.

PA (ORTM/) ORTMANN R.

PA (LUEO/) LUEOEND R.

PA (STAU/) STAUFENBIEL M.

PA (FREY/) FREY P.

XX

PI Maurer P, Lechner F, Ortmann R, Lueoend R, Staufenbiel M, Frey P;

PI Renner WA, Bachmann M, Tissot A, Sebbel P, Piossek C;

XX

DR WPI; 2002-636514/68.

XX

PT Molecular antigen array used in the production of vaccines for infectious
PT diseases.

XX

PS Disclosure; Page 346-347; 418pp; English.

XX

CC The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold comprising: (i) a core particle selected from: (1) a
CC core particle of a non-natural origin; and (2) a core particle of natural
CC origin; and (ii) an organiser comprising at least one first attachment
CC site, where the organiser is connected to the core particle by at least
CC one covalent bond; (b) an antigen or antigenic determinant with at least
CC one second attachment site, where the antigen or antigenic determinant is
CC amyloid beta peptide (Abeta 1-42) or its fragment, and where the second
CC attachment site is selected from: (i) an attachment site not naturally
CC occurring with the antigen or antigenic determinant; and (ii) an
CC attachment site naturally occurring with the antigen or antigenic
CC determinant, where the second attachment site is capable of association
CC through at least one non-peptide bond to the first attachment site; and
CC where the antigen or antigenic determinant and the scaffold interact
CC through the association to form an ordered and repetitive antigen array.
CC Also included is a process for producing a non-naturally occurring
CC ordered and repetitive antigen array. The composition is used in

CC immunisation and as a vaccine for diseases such as influenza, graft
CC versus host disease, IgE-mediated allergic reactions, anaphylaxis, adult
CC respiratory distress syndrome (ARDS), Crohn's disease, allergic asthma,
CC acute lymphoblastic leukaemia, non-Hodgkin's lymphoma, Grave's disease,
CC systemic lupus erythematosus, inflammatory immune diseases, myasthenia
CC gravis, immunoproliferative disease lymphadenopathy,
CC angioimmunoproliferative lymphadenopathy, immunoblastic lymphadenopathy,
CC rheumatoid arthritis, diabetes, multiple sclerosis, Alzheimer's disease,
CC osteoporosis and infectious diseases. The antigens are modified to possess
CC a cleavage site (enterokinase or factor Xa) and a Cysteine- containing N-
CC or C-terminal linker peptide which serves as the attachment point to a
CC virus like particle or bacterial protein (the scaffold protein). The
CC present sequence is bacterial protein or peptide which is coupled to the
CC modified antigen to form the molecular antigen array

XX

SQ Sequence 212 AA;

Query Match 68.9%; Score 31; DB 5; Length 212;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYATEVXDxD 10
||: |: ||
Db 98 FYSWEIADKD 107

Search completed: February 10, 2005, 15:48:42
Job time : 88.1831 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 22.3944 Seconds
(without alignments)
33.334 Million cell updates/sec

Title: US-10-067-484-5
Perfect score: 45
Sequence: 1 FYATEVXDxD 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	34	75.6	346	4	US-09-940-244-379	Sequence 379, App
2	31	68.9	243	4	US-09-489-039A-8468	Sequence 8468, Ap
3	31	68.9	270	4	US-09-248-796A-14475	Sequence 14475, A
4	30	66.7	162	4	US-09-765-815-14	Sequence 14, Appl
5	30	66.7	433	1	US-08-417-492-2	Sequence 2, Appli
6	30	66.7	439	4	US-09-489-039A-8498	Sequence 8498, Ap
7	30	66.7	638	2	US-08-557-122A-38	Sequence 38, Appl
8	30	66.7	638	3	US-09-262-666-38	Sequence 38, Appl
9	30	66.7	698	4	US-09-949-016-10644	Sequence 10644, A
10	30	66.7	703	4	US-09-248-796A-14529	Sequence 14529, A
11	30	66.7	1027	4	US-09-107-532A-6675	Sequence 6675, Ap
12	30	66.7	1072	4	US-09-902-540-15572	Sequence 15572, A
13	30	66.7	1371	4	US-09-902-540-16024	Sequence 16024, A
14	30	66.7	2680	4	US-09-489-039A-7973	Sequence 7973, Ap
15	29	64.4	149	4	US-09-270-767-40126	Sequence 40126, A
16	29	64.4	149	4	US-09-270-767-55342	Sequence 55342, A
17	29	64.4	149	4	US-09-471-276-1517	Sequence 1517, Ap
18	29	64.4	347	4	US-09-538-092-753	Sequence 753, App
19	29	64.4	362	4	US-09-634-238-417	Sequence 417, App
20	29	64.4	644	4	US-09-949-016-8212	Sequence 8212, Ap
21	29	64.4	865	4	US-09-902-540-10416	Sequence 10416, A
22	29	64.4	877	4	US-09-165-396-5	Sequence 5, Appli
23	28	62.2	139	4	US-09-909-650B-27	Sequence 27, Appl
24	28	62.2	221	4	US-09-902-540-16354	Sequence 16354, A
25	28	62.2	270	4	US-09-489-039A-14315	Sequence 14315, A
26	28	62.2	308	4	US-09-328-352-6762	Sequence 6762, Ap
27	28	62.2	322	4	US-09-134-000C-6420	Sequence 6420, Ap
28	28	62.2	339	4	US-09-583-110-3268	Sequence 3268, Ap
29	28	62.2	346	4	US-09-107-433-4133	Sequence 4133, Ap
30	28	62.2	375	2	US-08-837-593-5	Sequence 5, Appli
31	28	62.2	375	4	US-09-623-034-2	Sequence 2, Appli
32	28	62.2	384	4	US-09-909-650B-23	Sequence 23, Appl
33	28	62.2	393	4	US-09-393-858-2	Sequence 2, Appli
34	28	62.2	393	4	US-10-190-279-2	Sequence 2, Appli
35	28	62.2	422	3	US-09-025-580-3	Sequence 3, Appli
36	28	62.2	422	3	US-09-457-040B-38	Sequence 38, Appl
37	28	62.2	422	4	US-09-328-352-7923	Sequence 7923, Ap
38	28	62.2	424	4	US-09-909-650B-30	Sequence 30, Appl
39	28	62.2	426	4	US-09-909-650B-25	Sequence 25, Appl
40	28	62.2	434	1	US-07-952-817-9	Sequence 9, Appli

41	28	62.2	434	1	US-07-952-817-14	Sequence 14, Appl
42	28	62.2	434	6	5210025-2	Patent No. 5210025
43	28	62.2	434	6	5210025-7	Patent No. 5210025
44	28	62.2	434	6	5210025-2	Patent No. 5210025
45	28	62.2	434	6	5210025-7	Patent No. 5210025

ALIGNMENTS

RESULT 1

US-09-940-244-379
; Sequence 379, Application US/09940244
; Patent No. 6692917
; GENERAL INFORMATION:
; APPLICANT: Neri, Bruce P.
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor
; APPLICANT: Smith, Lloyd M.
; TITLE OF INVENTION: Reactions on Dendrimers
; FILE REFERENCE: FORS-06478
; CURRENT APPLICATION NUMBER: US/09/940,244
; CURRENT FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 379
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-940-244-379

Query Match 75.6%; Score 34; DB 4; Length 346;
Best Local Similarity 77.8%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
||||| |
Db 296 YATEVRDPD 304

RESULT 2

US-09-489-039A-8468
; Sequence 8468, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 8468
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8468

Query Match 68.9%; Score 31; DB 4; Length 243;
Best Local Similarity 60.0%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
|:| | | |
Db 97 FHATEAPDVD 106

RESULT 3

US-09-248-796A-14475
; Sequence 14475, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA
ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14475
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14475

Query Match 68.9%; Score 31; DB 4; Length 270;
Best Local Similarity 50.0%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
:| | | | :
Db 245 YYATETDDAE 254

RESULT 4

US-09-765-815-14
; Sequence 14, Application US/09765815
; Patent No. 6673586
; GENERAL INFORMATION:
; APPLICANT: Balk, Steven
; TITLE OF INVENTION: No. 6673586el Steroid Hormone Receptor
; TITLE OF INVENTION: Interacting Protein Kinase
; FILE REFERENCE: 01948/068002
; CURRENT APPLICATION NUMBER: US/09/765,815

; CURRENT FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/176,859
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-765-815-14

Query Match 66.7%; Score 30; DB 4; Length 162;
Best Local Similarity 85.7%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YATEVXD 8
|||
Db 104 YATEVVD 110

RESULT 5

US-08-417-492-2

; Sequence 2, Application US/08417492
; Patent No. 5750872
; GENERAL INFORMATION:
; APPLICANT: Bennett, Alan B
; APPLICANT: Brummell, David A
; APPLICANT: Grantz, Alexander A
; TITLE OF INVENTION: Nucleic Acids Encoding Ascorbate Free
; TITLE OF INVENTION: Radical Reductase and Their Uses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,492
; FILING DATE: 05-APR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-586US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-417-492-2

Query Match 66.7%; Score 30; DB 1; Length 433;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FYATEVXD XD 10
||| : |||
Db 143 FYLREIDDAD 152

RESULT 6

US-09-489-039A-8498
; Sequence 8498, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8498
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8498

Query Match 66.7%; Score 30; DB 4; Length 439;
Best Local Similarity 62.5%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8
|||| : |
Db 245 FYATAISD 252

RESULT 7

US-08-557-122A-38
; Sequence 38, Application US/08557122A
; Patent No. 5879664
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Mailand
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5879664o No. 5879664disk of No. 5879664th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York


```

; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,122A
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-557-122A-38

```

```

Query Match          66.7%; Score 30; DB 2; Length 638;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches      6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 YATEVXD 8
        |||||
Db     451 YATEVKD 457

```

RESULT 8

US-09-262-666-38

; Sequence 38, Application US/09262666

; Patent No. 6346244

; GENERAL INFORMATION:

; APPLICANT: Hjort, Carsten Mailand

; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6346244o No. 6346244disk of No. 6346244th America, Inc.

; STREET: 405 Lexington Avenue, 64th Floor

; CITY: New York

; STATE: New York

; COUNTRY: United States of America

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/09/262,666
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,122
; FILING DATE: 11-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-262-666-38

```

```

Query Match          66.7%; Score 30; DB 3; Length 638;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches      6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      2 YATEVXD 8
        |||||
Db      451 YATEVKD 457

```

RESULT 9

```

US-09-949-016-10644
; Sequence 10644, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES
THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10644
; LENGTH: 698
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10644

```

Query Match 66.7%; Score 30; DB 4; Length 698;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FYATEV 6
|||||:
Db 306 FYATEI 311

RESULT 10

US-09-248-796A-14529
; Sequence 14529, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA
ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14529
; LENGTH: 703
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14529

Query Match 66.7%; Score 30; DB 4; Length 703;
Best Local Similarity 62.5%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FYATEVXD 8
|| ||: |
Db 134 FYPTIED 141

RESULT 11

US-09-107-532A-6675
; Sequence 6675, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND
THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA

```

;           ZIP: 02354
;   COMPUTER READABLE FORM:
;           MEDIUM TYPE: CD-ROM ISO9660
;           COMPUTER: PC
;           OPERATING SYSTEM: <Unknown>
;           SOFTWARE: ASCII
;   CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/107,532A
;           FILING DATE: 30-Jun-1998
;   PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 60/085,598
;           FILING DATE: 14 May 1998
;           APPLICATION NUMBER: 60/051571
;           FILING DATE: July 2, 1997
;   ATTORNEY/AGENT INFORMATION:
;           NAME: Ariniello, Pamela Deneke
;           REGISTRATION NUMBER: 40,489
;           REFERENCE/DOCKET NUMBER: GTC-012
;   TELECOMMUNICATION INFORMATION:
;           TELEPHONE: (781)893-5007
;           TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 6675:
;   SEQUENCE CHARACTERISTICS:
;           LENGTH: 1027 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: YES
;   ORIGINAL SOURCE:
;           ORGANISM: Enterococcus faecium
;   FEATURE:
;           NAME/KEY: misc_feature
;           LOCATION: (B) LOCATION 1...1027
;   SEQUENCE DESCRIPTION: SEQ ID NO: 6675:
US-09-107-532A-6675

```

```

Query Match          66.7%;  Score 30;  DB 4;  Length 1027;
Best Local Similarity 66.7%;  Pred. No. 6.5e+02;
Matches      6;  Conservative    0;  Mismatches    3;  Indels      0;  Gaps      0;

```

```

Qy      2 YATEVXDxD 10
        || || ||
Db      453 YALEVTDVD 461

```

RESULT 12

```

US-09-902-540-15572
; Sequence 15572, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;   APPLICANT: Goldman, Barry S.
;   APPLICANT: Hinkle, Gregory J.
;   APPLICANT: Slater, Steven C.
;   APPLICANT: Wiegand, Roger C.
;   TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;   FILE REFERENCE: 38-10(15849)B
;   CURRENT APPLICATION NUMBER: US/09/902,540

```

; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15572
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15572

Query Match 66.7%; Score 30; DB 4; Length 1072;
Best Local Similarity 55.6%; Pred. No. 6.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
|||::||
Db 992 YATDLFDAD 1000

RESULT 13

US-09-902-540-16024
; Sequence 16024, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16024
; LENGTH: 1371
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16024

Query Match 66.7%; Score 30; DB 4; Length 1371;
Best Local Similarity 55.6%; Pred. No. 9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
|||::||
Db 862 YATDLFDAD 870

RESULT 14

US-09-489-039A-7973
; Sequence 7973, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 7973
 ; LENGTH: 2680
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-7973

Query Match 66.7%; Score 30; DB 4; Length 2680;
 Best Local Similarity 62.5%; Pred. No. 1.9e+03;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8
 |||:|
 Db 2123 FYVTDVTD 2130

RESULT 15

US-09-270-767-40126
 ; Sequence 40126, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 40126
 ; LENGTH: 149
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 US-09-270-767-40126

Query Match 64.4%; Score 29; DB 4; Length 149;
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXD 10
 :||:|
 Db 68 YYASEVQSAD 77

Search completed: February 10, 2005, 16:02:08
 Job time : 22.3944 secs

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:49:10 ; Search time 59.8592 Seconds
(without alignments)
54.586 Million cell updates/sec

Title: US-10-067-484-5
Perfect score: 45
Sequence: 1 FYATEVXDxD 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					
Result		Query					Description
No.	Score	Match	Length	DB	ID		
1	41	91.1	10	14	US-10-067-484-5		Sequence 5, Appli

2	41	91.1	10	14	US-10-067-620-5	Sequence 5, Appli
3	35	77.8	13	14	US-10-067-484-7	Sequence 7, Appli
4	35	77.8	13	14	US-10-067-620-7	Sequence 7, Appli
5	34	75.6	346	10	US-09-940-244-379	Sequence 379, App
6	34	75.6	398	10	US-09-769-787-61	Sequence 61, Appl
7	34	75.6	398	17	US-10-472-928-3644	Sequence 3644, Ap
8	33	73.3	544	15	US-10-424-599-148458	Sequence 148458,
9	33	73.3	558	15	US-10-425-114-57793	Sequence 57793, A
10	32	71.1	865	16	US-10-771-931-43	Sequence 43, Appl
11	32	71.1	975	15	US-10-369-493-13146	Sequence 13146, A
12	31	68.9	92	11	US-09-864-408A-1260	Sequence 1260, Ap
13	31	68.9	170	16	US-10-767-701-32427	Sequence 32427, A
14	31	68.9	212	10	US-09-848-616-139	Sequence 139, App
15	31	68.9	212	14	US-10-289-454-1	Sequence 1, Appli
16	31	68.9	212	14	US-10-050-902-139	Sequence 139, App
17	31	68.9	212	14	US-10-050-898-139	Sequence 139, App
18	31	68.9	212	15	US-10-346-190-1	Sequence 1, Appli
19	31	68.9	212	15	US-10-289-456-1	Sequence 1, Appli
20	31	68.9	381	15	US-10-282-122A-50829	Sequence 50829, A
21	31	68.9	401	16	US-10-437-963-173368	Sequence 173368,
22	31	68.9	408	14	US-10-156-761-14852	Sequence 14852, A
23	31	68.9	433	15	US-10-424-599-230372	Sequence 230372,
24	31	68.9	433	15	US-10-424-599-243063	Sequence 243063,
25	31	68.9	452	15	US-10-425-114-36611	Sequence 36611, A
26	31	68.9	492	15	US-10-424-599-225164	Sequence 225164,
27	31	68.9	500	17	US-10-866-527-38	Sequence 38, Appl
28	31	68.9	516	15	US-10-282-122A-53219	Sequence 53219, A
29	31	68.9	524	16	US-10-437-963-146419	Sequence 146419,
30	30	66.7	44	15	US-10-431-596-70	Sequence 70, Appl
31	30	66.7	52	15	US-10-424-599-182586	Sequence 182586,
32	30	66.7	162	10	US-09-765-815-14	Sequence 14, Appl
33	30	66.7	229	9	US-09-815-242-12391	Sequence 12391, A
34	30	66.7	229	15	US-10-282-122A-44014	Sequence 44014, A
35	30	66.7	238	15	US-10-425-114-38222	Sequence 38222, A
36	30	66.7	277	14	US-10-017-161-2062	Sequence 2062, Ap
37	30	66.7	277	15	US-10-292-798-1708	Sequence 1708, Ap
38	30	66.7	327	14	US-10-156-761-9826	Sequence 9826, Ap
39	30	66.7	408	15	US-10-282-122A-53311	Sequence 53311, A
40	30	66.7	446	10	US-09-955-999-67	Sequence 67, Appl
41	30	66.7	485	17	US-10-866-527-48	Sequence 48, Appl
42	30	66.7	508	15	US-10-282-122A-52721	Sequence 52721, A
43	30	66.7	514	16	US-10-437-963-162769	Sequence 162769,
44	30	66.7	537	15	US-10-425-114-64801	Sequence 64801, A
45	30	66.7	537	15	US-10-425-114-67015	Sequence 67015, A

ALIGNMENTS

RESULT 1

US-10-067-484-5

; Sequence 5, Application US/10067484

; Publication No. US20030170763A1

; GENERAL INFORMATION:

; APPLICANT: Buchanan, Bob B.

; APPLICANT: del Val, Gregorio

; APPLICANT: Frick, Oscar L.


```
; TITLE OF INVENTION: RAGWEED ALLERGENS
; FILE REFERENCE: 416272000200
; CURRENT APPLICATION NUMBER: US/10/067,484
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Ragweed
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 7,9
;   OTHER INFORMATION: Xaa= Leucine or Isoleucine
US-10-067-484-5
```

```
Query Match          91.1%;  Score 41;  DB 14;  Length 10;
Best Local Similarity 100.0%;  Pred. No. 0.083;
Matches    10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 FYATEVXDxD 10
        |||||
Db      1 FYATEVXDxD 10
```

RESULT 2

```
US-10-067-620-5
; Sequence 5, Application US/10067620
; Publication No. US20030180225A1
; GENERAL INFORMATION:
; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; APPLICANT: Teuber, Suzanne S.
; TITLE OF INVENTION: WALNUT AND RYEGRASS ALLERGENS
; FILE REFERENCE: 416272003400
; CURRENT APPLICATION NUMBER: US/10/067,620
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Ragweed
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 7,9
;   OTHER INFORMATION: Xaa= Leucine or Isoleucine
US-10-067-620-5
```

```
Query Match          91.1%;  Score 41;  DB 14;  Length 10;
Best Local Similarity 100.0%;  Pred. No. 0.083;
Matches    10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

Qy 1 FYATEVXDxD 10
| | | | | | | |
Db 1 FYATEVXDxD 10

RESULT 3

US-10-067-484-7
; Sequence 7, Application US/10067484
; Publication No. US20030170763A1
; GENERAL INFORMATION:
; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; TITLE OF INVENTION: RAGWEED ALLERGENS
; FILE REFERENCE: 416272000200
; CURRENT APPLICATION NUMBER: US/10/067,484
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Ragweed
US-10-067-484-7

Query Match 77.8%; Score 35; DB 14; Length 13;
Best Local Similarity 77.8%; Pred. No. 1.8;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
| | | | | | | |
Db 2 YATEVLDLD 10

RESULT 4

US-10-067-620-7
; Sequence 7, Application US/10067620
; Publication No. US20030180225A1
; GENERAL INFORMATION:
; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; APPLICANT: Teuber, Suzanne S.
; TITLE OF INVENTION: WALNUT AND RYEGRASS ALLERGENS
; FILE REFERENCE: 416272003400
; CURRENT APPLICATION NUMBER: US/10/067,620
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 13
; TYPE: PRT

; ORGANISM: Ragweed
US-10-067-620-7

Query Match 77.8%; Score 35; DB 14; Length 13;
Best Local Similarity 77.8%; Pred. No. 1.8;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
 ||||| |
Db 2 YATEVLdLD 10

RESULT 5

US-09-940-244-379
; Sequence 379, Application US/09940244
; Publication No. US20030044796A1
; GENERAL INFORMATION:
; APPLICANT: Neri, Bruce P.
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor
; APPLICANT: Smith, Lloyd M.
; TITLE OF INVENTION: Reactions on Dendrimers
; FILE REFERENCE: FORS-06478
; CURRENT APPLICATION NUMBER: US/09/940,244
; CURRENT FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 379
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-940-244-379

Query Match 75.6%; Score 34; DB 10; Length 346;
Best Local Similarity 77.8%; Pred. No. 87;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
 ||||| |
Db 296 YATEVRDPD 304

RESULT 6

US-09-769-787-61
; Sequence 61, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1

; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-61

Query Match 75.6%; Score 34; DB 10; Length 398;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
|:||||:|
Db 86 FFATEVVESD 95

RESULT 7

US-10-472-928-3644
; Sequence 3644, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3644
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: glycosyl transferase, family 8
; OTHER INFORMATION: Cellular location: cytoplasm
; OTHER INFORMATION: Similar to strain R6 sequence 15902441 (8.E-31)
US-10-472-928-3644

Query Match 75.6%; Score 34; DB 17; Length 398;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
|:||||:|
Db 86 FFATEVVESD 95

RESULT 8

US-10-424-599-148458
; Sequence 148458, Application US/10424599

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 148458
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(544)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10507C.1.pep
US-10-424-599-148458

Query Match 73.3%; Score 33; DB 15; Length 544;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
|| ||| : |
Db 222 FYCTEVSND 231

RESULT 9

US-10-425-114-57793
; Sequence 57793, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57793
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:

; OTHER INFORMATION: Clone ID: UC-GMROPIC109C08_FLI.pep
US-10-425-114-57793

Query Match 73.3%; Score 33; DB 15; Length 558;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
|| ||| : |
Db 235 FYCTEVSND 244

RESULT 10

US-10-771-931-43

; Sequence 43, Application US/10771931
; Publication No. US20040151737A1
; GENERAL INFORMATION:
; APPLICANT: Courtney, Harry
; TITLE OF INVENTION: Streptococcal Serum Opacity Factors And Fibronectin-
Binding Proteins And
; TITLE OF INVENTION: Peptides Thereof For The Treatment And Detection of
Streptococcal Infection
; FILE REFERENCE: 13314.1001U
; CURRENT APPLICATION NUMBER: US/10/771,931
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-771-931-43

Query Match 71.1%; Score 32; DB 16; Length 865;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
||: ||| | |
Db 544 FYSVDVTDS 553

RESULT 11

US-10-369-493-13146

; Sequence 13146, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION
OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13146
; LENGTH: 975
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(975)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13146

Query Match 71.1%; Score 32; DB 15; Length 975;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YATEVXDxD 10
| | | | |
Db 933 YVTEVSDDL 941

RESULT 12

US-09-864-408A-1260
; Sequence 1260, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and
Polypeptides Encoded Thereby
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1260
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-1260

Query Match 68.9%; Score 31; DB 11; Length 92;
Best Local Similarity 66.7%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YATEVXDxD 10
| | | | |
Db 56 YVTEVLDDD 64

RESULT 13

US-10-767-701-32427
; Sequence 32427, Application US/10767701

; Publication No. US20040172684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
 With
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 ; FILE REFERENCE: 38-21(53535)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-29
 ; NUMBER OF SEQ ID NOS: 63128
 ; SEQ ID NO 32427
 ; LENGTH: 170
 ; TYPE: PRT
 ; ORGANISM: Sorghum bicolor
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C12971_1.pep
 US-10-767-701-32427

Query Match 68.9%; Score 31; DB 16; Length 170;
 Best Local Similarity 55.6%; Pred. No. 1.7e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
 |||::||
 Db 110 YATQIVDL 118

RESULT 14

US-09-848-616-139
 ; Sequence 139, Application US/09848616
 ; Publication No. US20030054010A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sebbel, Peter
 ; APPLICANT: Dunant, Nicolas
 ; APPLICANT: Bachmann, Martin
 ; APPLICANT: Tissot, Alain
 ; APPLICANT: Lechner, Franziska
 ; TITLE OF INVENTION: Molecular Antigen Array
 ; FILE REFERENCE: 1700.0180002
 ; CURRENT APPLICATION NUMBER: US/09/848,616
 ; CURRENT FILING DATE: 2001-05-05
 ; NUMBER OF SEQ ID NOS: 186
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 139
 ; LENGTH: 212
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-09-848-616-139

Query Match 68.9%; Score 31; DB 10; Length 212;
 Best Local Similarity 50.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
 ||:|:|

Db 98 FYSWEIADKD 107

RESULT 15

US-10-289-454-1

; Sequence 1, Application US/10289454

; Publication No. US20030157479A1

; GENERAL INFORMATION:

; APPLICANT: Bachmann, Martin

; APPLICANT: Jennings, Gary

; APPLICANT: Sonderegger, Ivo

; TITLE OF INVENTION: Antigen Arrays for Treatments of Allergic Eosinophilic Diseases

; FILE REFERENCE: 1700.0360001

; CURRENT APPLICATION NUMBER: US/10/289,454

; CURRENT FILING DATE: 2003-02-10

; PRIOR APPLICATION NUMBER: US 60/396,636

; PRIOR FILING DATE: 2002-07-19

; PRIOR APPLICATION NUMBER: PCT/IB02/00166

; PRIOR FILING DATE: 2002-01-21

; PRIOR APPLICATION NUMBER: US 10/050,902

; PRIOR FILING DATE: 2002-01-18

; PRIOR APPLICATION NUMBER: US 60/331,045

; PRIOR FILING DATE: 2001-11-07

; NUMBER OF SEQ ID NOS: 386

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 212

; TYPE: PRT

; ORGANISM: Haemophilus influenzae

US-10-239-454-1

Query Match 68.9%; Score 31; DB 14; Length 212;

Best Local Similarity 50.0%; Pred. No. 2.1e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10

||: |: ||

Db 98 FYSWEIADKD 107

Search completed: February 10, 2005, 16:41:32

Job time : 60.8592 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 15.493 Seconds

(without alignments)

62.104 Million cell updates/sec

Title: US-10-067-484-5

Perfect score: 45

Sequence: 1 FYATEVXDxD 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	34	75.6	398	2	G95205	glycosyl transfera
2	32	71.1	357	2	B81292	hypothetical prote
3	32	71.1	510	2	B60280	bacillolysin-like
4	32	71.1	829	2	E87305	TonB-dependent rec
5	32	71.1	1186	2	S70430	hypothetical prote
6	31	68.9	63	2	E69800	hypothetical prote
7	31	68.9	210	2	E75315	probable c-type cy
8	31	68.9	212	2	C43310	stringent starvati
9	31	68.9	418	2	E65014	xanthosine permeas
10	31	68.9	428	2	T06464	protein kinase (EC
11	31	68.9	433	2	A55333	monodehydroascorba
12	31	68.9	464	2	T03780	probable integral
13	31	68.9	710	2	T26742	hypothetical prote
14	31	68.9	1079	2	T18356	membrane protein p
15	31	68.9	1207	2	B88789	protein ZK1251.9 [
16	31	68.9	1211	2	T23210	hypothetical prote
17	30	66.7	114	2	S77061	transposase sll066
18	30	66.7	176	2	AB0777	probable lipoprote
19	30	66.7	229	2	D90002	hypothetical prote
20	30	66.7	257	2	T34089	hypothetical prote
21	30	66.7	259	2	S76643	transposase slr051
22	30	66.7	261	2	S75081	transposase slr026
23	30	66.7	261	2	S77171	transposase sll171
24	30	66.7	261	2	S77351	transposase sll171
25	30	66.7	261	2	S76309	transposase slr035
26	30	66.7	305	2	C69465	dinitrogenase redu
27	30	66.7	314	1	WMBEB4	ribonucleoside-dip

28	30	66.7	314	2	H88991	protein K08D9.1 [i
29	30	66.7	433	2	T06407	monodehydroascorba
30	30	66.7	434	2	JU0182	monodehydroascorba
31	30	66.7	571	2	AG0144	D-lactate dehydrog
32	30	66.7	584	2	S06318	endoplasmic reticu
33	30	66.7	638	1	ISMSE	protein disulfide-
34	30	66.7	643	1	S32476	protein disulfide-
35	30	66.7	688	1	JC1469	beta-adrenergic-re
36	30	66.7	688	1	A39336	beta-adrenergic-re
37	30	66.7	882	2	E96931	hypothetical prote
38	30	66.7	1010	2	T36383	probable large ATP
39	29	64.4	115	2	H72643	hypothetical prote
40	29	64.4	230	2	B86824	two-component syst
41	29	64.4	256	2	AI1204	molybdate ABC tran
42	29	64.4	326	2	T09259	cathepsin L-like p
43	29	64.4	326	2	E84812	hypothetical prote
44	29	64.4	347	2	S67159	probable membrane
45	29	64.4	388	2	D84992	hypothetical prote

ALIGNMENTS

RESULT 1

G95205

glycosyl transferase, family 8 SP1765 [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C;Accession: G95205

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidelberg, J.; DeBoy, R.T.; Haft, D.H.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.; Kolonay, J.F.; Nelson, W.C.; Peterson, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, E.; Khouri, H.; Wolf, A.M.; Utterback, T.R.; Hansen, C.L.; McDonald, L.A.; Feldblyum, T.V.; Angiuoli, S.; Dickinson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, D.A.; Hollingshead, S.K.; Fraser, C.M.

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: G95205

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-398 <KUR>

A;Cross-references: UNIPROT:Q97P77; GB:AE005672; PIDN:AAK75840.1; PID:g14973262; GSPDB:GN00164; TIGR:SP4SP1765

A;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP1765

Query Match 75.6%; Score 34; DB 2; Length 398;
 Best Local Similarity 60.0%; Pred. No. 9.8;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10

Db |:||||:|
86 FFATEVVESD 95

RESULT 2

B81292

hypothetical protein Cj1459 [imported] - Campylobacter jejuni (strain NCTC 11168)

C;Species: Campylobacter jejuni

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C;Accession: B81292

R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R.M.; Feltwell, T.; Holroyd, S.; Jagels, K.; Karlyshev, A.; Moule, S.; Pallen, M.J.; Penn, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.G.

Nature 403, 665-668, 2000

A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences.

A;Reference number: A81250; MUID:20150912; PMID:10688204

A;Accession: B81292

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-357 <PAR>

A;Cross-references: UNIPROT:Q9PMK1; GB:AL139078; GB:AL111168; NID:g6969723;

PIDN:CAB73882.1; PID:g6968887; GSPDB:GN00120; CJSP:Cj1459

A;Experimental source: serotype O2, strain NCTC 11168

C;Genetics:

A;Gene: Cj1459

Query Match 71.1%; Score 32; DB 2; Length 357;
Best Local Similarity 62.5%; Pred. No. 25;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8

:||||:|

Db 130 YYATEILD 137

RESULT 3

B60280

bacillolysin-like proteinase (EC 3.4.24.-) prtA precursor - Listeria monocytogenes (strain LM8, serotype 4b)

N;Alternate names: metalloproteinase homolog mpl

C;Species: Listeria monocytogenes

C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 09-Jul-2004

C;Accession: B60280; A43868; S24232

R;Mengaud, J.; Geoffroy, C.; Cossart, P.

Infect. Immun. 59, 1043-1049, 1991

A;Title: Identification of a new operon involved in Listeria monocytogenes virulence: its first gene encodes a protein homologous to bacterial metalloproteases.

A;Reference number: A60280; MUID:91147180; PMID:1705239

A;Accession: B60280

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-510 <MEN>

A;Cross-references: UNIPROT:P34025

A;Experimental source: strain LM8, serotype 4b
R;Vazquez-Boland, J.A.; Kocks, C.; Dramsi, S.; Ohayon, H.; Geoffroy, C.; Mengaud, J.; Cossart, P.
Infect. Immun. 60, 219-230, 1992
A;Title: Nucleotide sequence of the lecithinase operon of *Listeria monocytogenes* and possible role of lecithinase in cell-to-cell spread.
A;Reference number: A43868; MUID:92104678; PMID:1309513
A;Accession: A43868
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 504-510 <VAZ>
A;Note: sequence extracted from NCBI backbone (NCBIN:74437, NCBIP:74457)
R;Rasmussen, O.F.; Beck, T.; Olsen, J.E.; Dons, L.; Rossen, L.
Infect. Immun. 59, 3945-3951, 1991
A;Title: *Listeria monocytogenes* isolates can be classified into two major types according to the sequence of the listeriolysin gene.
A;Reference number: S24230; MUID:92040062; PMID:1937753
A;Accession: S24232
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-46, 'A', 48-102, 'A', 104-271 <RAS>
A;Cross-references: EMBL:X60035
A;Experimental source: strain 12067, serotype4b
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
C;Genetics:
A;Gene: prtA; mpl
C;Superfamily: zinc metalloendopeptidase, neutral protease type (elastase)
C;Keywords: extracellular protein; hydrolase; metalloproteinase; zinc
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-510/Product: bacillolysin-like proteinase #status predicted <MAT>

Query Match 71.1%; Score 32; DB 2; Length 510;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8
|||:|
Db 282 FYASEVYD 289

RESULT 4

E87305

TonB-dependent receptor [imported] - *Caulobacter crescentus*

C;Species: *Caulobacter crescentus*

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: E87305

R;Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.; Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.; Craven, M.; Khouri, H.; Shetty, J.; Berry, K.; Utterback, T.; Tran, K.; Wolf, A.; Vamathevan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: E87305
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-829 <STO>
A;Cross-references: UNIPROT:Q9AAY8; GB:AE005673; NID:g13421627; PIDN:AAK22441.1;
GSPDB:GN00148
C;Genetics:
A;Gene: CC0454

Query Match 71.1%; Score 32; DB 2; Length 829;
Best Local Similarity 60.0%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
|::||| | |
Db 335 FHSTEVFDPD 344

RESULT 5

S70430
hypothetical protein 4 - fruit fly (*Drosophila melanogaster*) retrotransposon
mgd1
C;Species: *Drosophila melanogaster*
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 07-May-1999
C;Accession: S70430
R;Avedisov, S.N.; Cherkasova, V.A.; Ilyin, Y.V.
Genetika 26, 1905-1914, 1990
A;Title: The primary structure features of the full-length copy of mgd1
Drosophila retrotransposon.
A;Reference number: S70427; MUID:91160952; PMID:1963611
A;Accession: S70430
A;Molecule type: DNA
A;Residues: 1-1186 <AVE>
A;Cross-references: EMBL:X59545
C;Genetics:
A;Gene: FlyBase:mgd1
A;Cross-references: FlyBase:FBgn0002697
A;Mobile element: retrotransposon mgd1
C;Superfamily: pol polyprotein

Query Match 71.1%; Score 32; DB 2; Length 1186;
Best Local Similarity 50.0%; Pred. No. 91;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
||: |: | |
Db 768 FYSNEIIDLD 777

RESULT 6

E69800
hypothetical protein yfhD - *Bacillus subtilis*
C;Species: *Bacillus subtilis*
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: E69800
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.;

Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.; Brüschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Guiseppi, G.; Guy, B.J.; Haga, K.; Haiech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott, A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.; Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenegger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: E69800

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-63 <KUN>

A;Cross-references: UNIPROT:O31572; GB:Z99108; GB:AL009126; NID:g2633055;

PIDN:CAB12678.1; PID:e1182839; PID:g2633173

A;Experimental source: strain 168

C;Genetics:

A;Gene: yfhD

Query Match 68.9%; Score 31; DB 2; Length 63;

Best Local Similarity 55.6%; Pred. No. 6.3;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10

|:|:|:|

Db 35 YSTELADAD 43

RESULT 7

E75315

probable c-type cytochrome - *Deinococcus radiodurans* (strain R1)

C;Species: *Deinococcus radiodurans*

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: E75315

R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.;
Dodson, R.J.; Haft, D.H.; Gwinn, M.L.; Nelson, W.C.; Richardson, D.L.; Moffat,
K.S.; Qin, H.; Jiang, L.; Pamphile, W.; Crosby, M.; Shen, M.; Vamathevan, J.J.;
Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Makarova, K.S.; Aravind, L.;
Daly, M.J.; Minton, K.W.; Fleischmann, R.D.; Ketchum, K.A.; Nelson, K.E.;
Salzberg, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans*
R1.

A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: E75315

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-210 <WHI>

A;Cross-references: UNIPROT:Q9RSM9; GB:AE002045; GB:AE000513; NID:g6459886;
PIDN:AAF11644.1; PID:g6459890; TIGR:DR2095; GSPDB:GN00077

A;Experimental source: strain R1

C;Genetics:

A;Gene: DR2095

A;Map position: 1

Query Match 68.9%; Score 31; DB 2; Length 210;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
| | | | |
Db 190 FVATQVSDQD 199

RESULT 8

C43310

stringent starvation protein homolog - *Haemophilus somnus*

C;Species: *Haemophilus somnus*

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 12-Jul-2004

C;Accession: C43310

R;Theisen, M.; Potter, A.A.

J. Bacteriol. 174, 17-23, 1992

A;Title: Cloning, sequencing, expression, and functional studies of a 15,000-
molecular-weight *Haemophilus somnus* antigen similar to *Escherichia coli*
ribosomal protein S9.

A;Reference number: A43310; MUID:92104958; PMID:1729207

A;Accession: C43310

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-212 <THE>

A;Cross-references: UNIPROT:P31784; GB:S75161; NID:g241865; PIDN:AAB20822.1;
PID:g241868

A;Experimental source: HS25

A;Note: sequence extracted from NCBI backbone (NCBIN:75161, NCBIP:75172)

C;Superfamily: stringent starvation protein A

Query Match 68.9%; Score 31; DB 2; Length 212;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10

||||: | :
Db 36 YATEIVDSE 44

RESULT 9

E65014

xanthosine permease - *Escherichia coli* (strain K-12)

C;Species: *Escherichia coli*

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C;Accession: E65014

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of *Escherichia coli* K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: E65014

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-418 <BLAT>

A;Cross-references: UNIPROT:P45562; GB:AE000328; GB:U00096; NID:g2367135;

PIDN:AAC75459.1; PID:g1788745; UWGP:b2406

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: xapB

Query Match 68.9%; Score 31; DB 2; Length 418;
Best Local Similarity 60.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
||| |||
Db 84 FYAASVTDPD 93

RESULT 10

T06464

protein kinase (EC 2.7.1.-) - garden pea

C;Species: *Pisum sativum* (garden pea)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 16-Aug-2004

C;Accession: T06464

R;Lin, X.; Watson, J.C.

Plant Physiol. 100, 1072-1074, 1993

A;Title: CDNA sequence of PspK5, a protein kinase homolog from *Pisum sativum* L.

A;Reference number: Z15698

A;Accession: T06464

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-428 <LIN>

A;Cross-references: UNIPROT:Q07526; EMBL:M92989; NID:g556346; PIDN:AAA50304.1;

PID:g556347

A;Experimental source: cv. Alaska

C;Genetics:

A;Gene: PK5

C;Superfamily: protein kinase homology

C;Keywords: phosphotransferase

F;101-387/Domain: protein kinase homology <KIN>

Query Match 68.9%; Score 31; DB 2; Length 428;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FYATEV 6
|||||
Db 207 FYATEV 212

RESULT 11

A55333

monodehydroascorbate reductase (NADH2) (EC 1.6.5.4) - garden pea

C;Species: Pisum sativum (garden pea)

C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004

C;Accession: A55333

R;Murthy, S.S.; Zilinskas, B.A.

J. Biol. Chem. 269, 31129-31133, 1994

A;Title: Molecular cloning and characterization of a cDNA encoding pea monodehydroascorbate reductase.

A;Reference number: A55333; MUID:95074153; PMID:7983054

A;Accession: A55333

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-433 <MUR>

A;Cross-references: UNIPROT:Q40977; GB:U06461; NID:g497119; PIDN:AAA60979.1; PID:g497120

C;Superfamily: rubredoxin-NAD+ reductase rubB

C;Keywords: NAD; oxidoreductase

Query Match 68.9%; Score 31; DB 2; Length 433;
Best Local Similarity 60.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
|| ||| |
Db 142 FYLREVDDAD 151

RESULT 12

T03780

probable integral membrane protein - rice

C;Species: Oryza sativa (rice)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T03780

R;Belouchi, A.; Kwan, T.; Gros, P.

Plant Mol. Biol. 33, 1085-1092, 1997

A;Title: Cloning and characterization of the OsNramp family from Oryza sativa, a new family of membrane proteins possibly implicated in the transport of metal ions.

A;Reference number: Z15079; MUID:97299840; PMID:9154989

A;Accession: T03780

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-464 <BEL>

A;Cross-references: UNIPROT:O24209; EMBL:L81152; NID:g2231164; PIDN:AAB61961.1; PID:g2231149

C;Genetics:

A;Gene: Nramp2

C;Superfamily: natural resistance-associated macrophage protein 1

Query Match 68.9%; Score 31; DB 2; Length 464;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FYATEV 6
|||||
Db 417 FYATEV 422

RESULT 13

T26742

hypothetical protein Y39A1A.22 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T26742

R;Wall, M.

submitted to the EMBL Data Library, September 1998

A;Reference number: Z20257

A;Accession: T26742

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-710 <WIL>

A;Cross-references: UNIPROT:Q9XX10; EMBL:AL031633; PIDN:CAA21031.1;

GSPDB:GN00021; CESP:Y39A1A.22

A;Experimental source: clone Y39A1A

C;Genetics:

A;Gene: CESP:Y39A1A.22

A;Map position: 3

A;Introns: 212/3

Query Match 68.9%; Score 31; DB 2; Length 710;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FYATEV 6
|||||
Db 412 FYATEV 417

RESULT 14

T18356

membrane protein p120 - Mycoplasma hominis

C;Species: Mycoplasma hominis

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T18356

R;Christiansen, G.; Mathiesen, S.L.; Nyvold, C.; Birkelund, S.

FEMS Microbiol. Lett. 121, 121-128, 1994

A;Title: Analysis of Mycoplasma hominis membrane protein, P120.

A;Reference number: Z18889; MUID:94364538; PMID:8082822

A;Accession: T18356

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1079 <CHR>

A;Cross-references: UNIPROT:Q49555; EMBL:X78450; NID:g587473; PID:g587474;
PIDN:CAA55207.1
C;Genetics:
A;Genetic code: SGC3
A;Note: P120

Query Match 68.9%; Score 31; DB 2; Length 1079;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXD 10
|:| || | |
Db 748 FFANEVDPDYD 757

RESULT 15

B88789.

protein ZK1251.9 [imported] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C;Accession: B88789

R;anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A;Title: Genome sequence of the nematode C. elegans: a platform for
investigating biology.

A;Reference number: A75000; MUID:99069613; PMID:9851916

A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and
www_sanger.ac.uk/Projects/C_elegans/ for a list of authors

A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103,
1999; and Science 285, 1493, 1999

A;Accession: B88789

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1207 <STO>

A;Cross-references: UNIPROT:Q21106; GB:chr_IV; PIDN:CAA92475.1; PID:g4008369;

GSPDB:GN00022; CESP:ZK1251.9

C;Genetics:

A;Gene: ZK1251.9

A;Map position: 4

Query Match 68.9%; Score 31; DB 2; Length 1207;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YATEVXD 8
||||| |
Db 812 YATEVTD 818

Search completed: February 10, 2005, 15:59:34
Job time : 24.493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 72.9577 Seconds
 (without alignments)
 70.188 Million cell updates/sec

Title: US-10-067-484-5
 Perfect score: 45
 Sequence: 1 FYATEVXDxD 10

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : UniProt_03:*
 1: uniprot_sprot:*
 2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	34	75.6	233	2	Q89RN8	Q89rn8 bradyrhizob
2	34	75.6	246	2	O40106	O40106 human immun
3	34	75.6	346	1	FEN_PYRAE	Q8zyn2 pyrobaculum
4	34	75.6	398	2	Q97P77	Q97p77 streptococc
5	34	75.6	449	2	Q88W47	Q88w47 lactobacill
6	34	75.6	946	2	Q7PQ81	Q7pq81 anopheles g
7	33	73.3	108	2	Q9J0Q9	Q9j0q9 human immun
8	33	73.3	410	2	Q8ZXA3	Q8zxa3 pyrobaculum
9	33	73.3	436	2	Q8A9A2	Q8a9a2 bacteroides
10	32	71.1	241	2	Q7X1P6	Q7x1p6 lactococcus
11	32	71.1	319	2	Q73P18	Q73p18 treponema d
12	32	71.1	357	2	Q9PMK1	Q9pmk1 campylobact
13	32	71.1	510	1	PRO2_LISMO	P34025 listeria mo
14	32	71.1	510	2	Q6E9N9	Q6e9n9 listeria mo
15	32	71.1	510	2	Q6EA99	Q6ea99 listeria mo
16	32	71.1	510	2	Q6EAC2	Q6eac2 listeria mo
17	32	71.1	510	2	Q6EAI1	Q6eai1 listeria mo
18	32	71.1	510	2	Q724L0	Q724l0 listeria mo
19	32	71.1	512	1	NRM4_ARATH	Q9fn18 arabidopsis
20	32	71.1	624	2	O97429	O97429 drosophila
21	32	71.1	809	2	Q72LY1	Q72ly1 leptospira

22	32	71.1	809	2	Q8EYF2	Q8eyf2 leptospira
23	32	71.1	829	2	Q9AAY8	Q9aay8 caulobacter
24	32	71.1	854	2	Q9S3P8	Q9s3p8 streptococc
25	32	71.1	862	2	Q9RPZ2	Q9rpz2 streptococc
26	32	71.1	865	2	Q9S4J9	Q9s4j9 streptococc
27	32	71.1	872	2	Q9S4J3	Q9s4j3 streptococc
28	32	71.1	1027	2	Q6A7Q8	Q6a7q8 propionibac
29	31	68.9	34	2	Q7SPS8	Q7sps8 human immun
30	31	68.9	34	2	Q7SPT1	Q7spt1 human immun
31	31	68.9	35	2	Q97632	Q97632 human immun
32	31	68.9	63	1	YFHD_BACSU	O31572 bacillus su
33	31	68.9	82	2	O56682	O56682 human immun
34	31	68.9	91	2	Q86879	Q86879 human immun
35	31	68.9	95	2	Q74529	Q74529 human immun
36	31	68.9	102	2	Q90RU5	Q90ru5 human immun
37	31	68.9	106	2	O37956	O37956 human immun
38	31	68.9	119	2	Q7ZKW7	Q7zkw7 human immun
39	31	68.9	134	2	Q7ZB10	Q7zb10 human immun
40	31	68.9	166	1	MDAF_CUCSA	P83966 cucumis sat
41	31	68.9	210	2	Q9RSM9	Q9rsm9 deinococcus
42	31	68.9	212	1	SSPA_HAESO	P31784 haemophilus
43	31	68.9	212	2	O50342	O50342 haemophilus
44	31	68.9	215	2	P88084	P88084 human immun
45	31	68.9	260	2	Q7MS87	Q7ms87 wolinnella s

ALIGNMENTS

RESULT 1

Q89RN8

ID Q89RN8 PRELIMINARY; PRT; 233 AA.
AC Q89RN8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Blr2725 protein.
GN OrderedLocusNames=blr2725;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005944; BAC47990.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000868; Iscrsm_hydrolase.

DR Pfam; PF00857; Isochorismatase; 1.
KW Complete proteome.
SQ SEQUENCE 233 AA; 25004 MW; 2AF118B6050504AF CRC64;

Query Match 75.6%; Score 34; DB 2; Length 233;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8
||||: |
Db 143 FYATELTD 150

RESULT 2

O40106

ID O40106 PRELIMINARY; PRT; 246 AA.
AC O40106;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
PX MEDLINE=97332315; PubMed=9188549;
RA Shioda T., Oka S., Xin X., Liu H., Harukuni R., Kurotani A.,
RA Fukushima M., Hasan M.K., Shiino T., Takebe Y., Iwamoto A., Nagai Y.;
RT "In vivo sequence variability of human immunodeficiency virus type 1
RT envelope gp120: association of V2 extention with slow disease
RT progression.";
RL J. Virol. 71:4871-4881(1997).
DR EMBL; AB002949; BAA21255.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 246 246
SQ SEQUENCE 246 AA; 27643 MW; 17A1B5D15AA428A0 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 246;
Best Local Similarity 70.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXD 10
|||| | |
Db 200 FYATEGIDGD 209

RESULT 3

FEN_PYRAE

ID FEN_PYRAE STANDARD; PRT; 346 AA.
 AC Q8ZYN2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Flap structure-specific endonuclease (EC 3.---).
 GN Name=fen; OrderedLocusNames=PAE0698;
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 RT aerophilum.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 CC -!- FUNCTION: Endonuclease that cleave the 5'overhanging flap
 CC structure that is generated by displacement synthesis when DNA
 CC polymerase encounters the 5'end of a downstream Okazaki fragment.
 CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
 CC Cleaves the junction between single and double-stranded regions of
 CC flap DNA (By similarity).
 CC -!- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
 CC -!- SIMILARITY: Belongs to the XPG/RAD2 endonuclease family. FEN1
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE009780; AAL62961.1; -.
 DR HSSP; O93634; 1B43.
 DR HAMAP; MF_00614; -, 1.
 DR InterPro; IPR008918; 5_3_exo_C.
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; Exo_N_I.
 DR InterPro; IPR006086; XPG_I.
 DR InterPro; IPR006085; XPG_N.
 DR InterPro; IPR006084; XPGC_Rad.
 DR Pfam; PF01367; 5_3_exonuc; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF00752; XPG_N; 1.
 DR PRINTS; PR00853; XPGGRADSUPER.
 DR SMART; SM00279; Hhh2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 DR PROSITE; PS00841; XPG_1; FALSE_NEG.
 KW Complete proteome; Endonuclease; Hydrolase; Magnesium; Metal-binding;
 KW Nuclease.

FT METAL 158 158 Magnesium 1 (By similarity).
SQ SEQUENCE 346 AA; 39099 MW; A9590463432AC1F7 CRC64;

Query Match 75.6%; Score 34; DB 1; Length 346;
Best Local Similarity 77.8%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVXDxD 10
||| | |
Db 296 YATEVRDPD 304

RESULT 4

Q97P77

ID Q97P77 PRELIMINARY; PRT; 398 AA.
AC Q97P77;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycosyl transferase, family 8.
GN OrderedLocusNames=SP1765;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL; AE007469; AAK75840.1; -.
DR PIR; G95205; G95205.
DR TIGR; SP1765; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0016758; F:transferase activity, transferring hexosyl . . .; IEA.
DR GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR InterPro; IPR002495; Glyco_trans_8.
DR Pfam; PF01501; Glyco_transf_8; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 398 AA; 46365 MW; 4404EF544488BB71 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 398;
Best Local Similarity 60.0%; Pred. No. 65;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
|: ||| : |

Db

86 FFATEVVESD 95

RESULT 5

Q88W47

ID Q88W47 PRELIMINARY; PRT; 449 AA.
AC Q88W47;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integral membrane protein.
GN OrderedLocusNames=lp_1815;
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFS1;
RX MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
DR EMBL; AL935257; CAD64224.1; -.
KW Complete proteome.
SQ SEQUENCE 449 AA; 50073 MW; 59A911D45E3E474B CRC64;

Query Match 75.6%; Score 34; DB 2; Length 449;
Best Local Similarity 77.8%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YATEVXDxD 10
| | | | | | |
Db 357 YATEVNDPD 365

RESULT 6

Q7PQ81

ID Q7PQ81 PRELIMINARY; PRT; 946 AA.
AC Q7PQ81;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP00000003976 (Fragment).
GN Name=ENSANGG00000003158;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;

RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAAB01008898; EAA09193.2; -.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR000357; HEAT.
 DR Pfam; PF02985; HEAT; 2.
 FT NON_TER 1 1
 SQ SEQUENCE 946 AA; 108703 MW; 7BBED36A20A4366F CRC64;

Query Match 75.6%; Score 34; DB 2; Length 946;
 Best Local Similarity 60.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXD XD 10
 : || || | |
 Db 935 YYAAEVNDAD 944

RESULT 7

Q9J0Q9

ID Q9J0Q9 PRELIMINARY; PRT; 108 AA.
 AC Q9J0Q9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN Name=eniv;
 OS Human immunodeficiency virus 1.
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21443129; PubMed=11559432; DOI=10.1089/088922201750461384;
 RA Reinis M., Bruckova M., Graham R.R., Vandasova J., Stankova M.,
 RA Carr J.K.;
 RT "Genetic subtypes of HIV type 1 viruses circulating in the Czech
 RT Republic.";
 RL AIDS Res. Hum. Retroviruses 17:1305-1310(2001).
 DR EMBL; AF223975; AAF34914.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 108 108
 SQ SEQUENCE 108 AA; 12067 MW; 6872EC962280B87B CRC64;

Query Match 73.3%; Score 33; DB 2; Length 108;
 Best Local Similarity 75.0%; Pred. No. 28;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8

||||:|
Db 57 FYATDVID 64

RESULT 8

Q8ZXA3

ID Q8ZXA3 PRELIMINARY; PRT; 410 AA.
AC Q8ZXA3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Acyl-CoA dehydrogenase.
GN OrderedLocusNames=PAE1378;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
PL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC -!- SIMILARITY: Belongs to the acyl-CoA dehydrogenase family.
DR EMBL; AE009818; AAL63446.1; -.
DR HSSP; Q06319; 1BUC.
DR GO; GO:0003995; F:acyl-CoA dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
KW Complete proteome; FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 410 AA; 46265 MW; B336E4CD930318CD CRC64;

Query Match 73.3%; Score 33; DB 2; Length 410;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYATEVXD 8
||||| :
Db 335 FYATEVAE 342

RESULT 9

Q8A9A2

ID Q8A9A2 PRELIMINARY; PRT; 436 AA.
AC Q8A9A2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative Fe-S oxidoreductase.
GN OrderedLocusNames=BT0913;
OS Bacteroides thetaiotaomicron.

OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
 RA Chiang H.C., Hooper L.V., Gordon J.I.;
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
 RL Science 299:2074-2076(2003).
 DR EMBL; AE016929; AAO76020.1; -.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0005506; F:iron ion binding; IEA.
 DR InterPro; IPR005840; Cons_hypoth1125.
 DR InterPro; IPR006638; Elp3/MiaB/NifB.
 DR InterPro; IPR007197; Radical_SAM.
 DR InterPro; IPR002792; TRAM.
 DR InterPro; IPR005839; UPF0004.
 DR Pfam; PF04055; Radical_SAM; 1.
 DR Pfam; PF00919; UPF0004; 1.
 DR SMART; SM00729; Elp3; 1.
 DR TIGRFAMs; TIGR01125; Cons_hypoth1125; 1.
 DR TIGRFAMs; TIGR00089; UPF0004; 1.
 DR PROSITE; PS50926; TRAM; 1.
 DR PROSITE; PS01278; UPF0004; 1.
 KW Complete proteome.
 SQ SEQUENCE 436 AA; 50869 MW; BC779FB6027574D4 CRC64;

Query Match 73.3%; Score 33; DB 2; Length 436;
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
 || || || ||
 Db 414 FYQVEVTDAD 423

RESULT 10

Q7X1P6

ID Q7X1P6 PRELIMINARY; PRT; 241 AA.
 AC Q7X1P6;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Lactococcus raffinolactis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1366;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 43920;
 RX MEDLINE=22338278; PubMed=12450840;
 RX DOI=10.1128/AEM.68.12.6152-6161.2002;
 RA Boucher I., Parrot M., Gaudreau H., Champagne C.P., Vadeboncoeur C.,
 RA Moineau S.;
 RT "Novel food-grade plasmid vector based on melibiose fermentation for

RT the genetic engineering of *Lactococcus lactis*.";
 RL Appl. Environ. Microbiol. 68:6152-6161(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 43920;
 RX MEDLINE=22723489; PubMed=12839781;
 RX DOI=10.1128/AEM.69.7.4049-4056.2003;
 RA Boucher I., Vadeboncoeur C., Moineau S.;
 RT "Characterization of genes involved in the metabolism of alpha-
 RT galactosides by *Lactococcus raffinolactis*.";
 RL Appl. Environ. Microbiol. 69:4049-4056(2003).
 DR EMBL; AY164273; AA026318.1; -.
 KW Hypothetical protein.
 FT NON_TER 241 241
 SQ SEQUENCE 241 AA; 27861 MW; 47B8C58052505027 CRC64;

Query Match 71.1%; Score 32; DB 2; Length 241;
 Best Local Similarity 60.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FYATEVXDxD 10
 ||||:| |
 Db 11 FYATQVQSDD 20

RESULT 11

Q73P18

ID Q73P18 PRELIMINARY; PRT; 319 AA.
 AC Q73P18;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=TDE0981;
 OS *Treponema denticola*.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; *Treponema*.
 OX NCBI_TaxID=158;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35405 / DSM 14222;
 RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
 RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
 RA Dodson R.J., Davidsen T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
 RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
 RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
 RA Gebregeorgis E., Geer K., Tsegaye G., Malek J.A., Ayodeji B.,
 RA Shatsman S., McLeod M.P., Smajs D., Howell J.K., Pal S., Amin A.,
 RA Vashisth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
 RA Weinstock G.M., Norris S.J., Fraser C.M., Paulsen I.T.;
 RT "Comparison of the genome of the oral pathogen *Treponema denticola*
 RT with other spirochete genomes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
 DR EMBL; AE017249; AAS11472.1; -.
 DR TIGR; TDE0981; -.
 KW Complete proteome.
 SQ SEQUENCE 319 AA; 36243 MW; EA224EDF0D7A4A9C CRC64;

Query Match 71.1%; Score 32; DB 2; Length 319;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYATEVXDxD 10
|||::|||
Db 290 FYQTKILDTD 299

RESULT 12

Q9PMK1

ID Q9PMK1 PRELIMINARY; PRT; 357 AA.
AC Q9PMK1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Cj1459.
GN OrderedLocusNames=Cj1459;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
RA Parkhill J., Wren B.W., Mungall K.L., Ketley J.M., Churcher C.M.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
DR EMBL; AL139078; CAB73882.1; -.
DR PIR; B81292; B81292.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 357 AA; 42358 MW; 03A81F58307082CF CRC64;

Query Match 71.1%; Score 32; DB 2; Length 357;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYATEVXD 8
:||||:|
Db 130 YYATEILD 137

RESULT 13

PRO2_LISMO

ID PRO2_LISMO STANDARD; PRT; 510 AA.
AC P34025;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Zinc metalloproteinase precursor (EC 3.4.24.-).
GN Name=mpl; Synonyms=prtA;

OS *Listeria monocytogenes*.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LO28 / Serovar 1/2c;
 RX MEDLINE=91147180; PubMed=1705239;
 RA Mengaud J., Geoffroy C., Cossart P.;
 RT "Identification of a new operon involved in *Listeria monocytogenes*
 RT virulence: its first gene encodes a protein homologous to bacterial
 RT metalloproteases.";
 RL Infect. Immun. 59:1043-1049(1991).
 RN [2]
 RP SEQUENCE OF 1-272 FROM N.A.
 RC STRAIN=12067;
 RX MEDLINE=92040062; PubMed=1937753;
 RA Rasmussen O.F., Beck T., Olsen J.E., Dons L., Rossen L.;
 RT "*Listeria monocytogenes* isolates can be classified into two major
 RT types according to the sequence of the listeriolysin gene.";
 RL Infect. Immun. 59:3945-3951(1991).
 CC -!- FUNCTION: Probably linked to the pathogenesis of listerial
 CC infection.
 CC -!- COFACTOR: Binds 1 zinc ion (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- INDUCTION: The mpl and the listeriolysin genes being physically
 CC linked, their expression may be regulated in a similar manner.
 CC -!- SIMILARITY: Belongs to the peptidase M4 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X60035; CAA42640.1; -.
 DR PIR; B60280; B60280.
 DR HSSP; P81177; 1BQB.
 DR MEROPS; M04.008; -.
 DR InterPro; IPR001570; Peptidase_M4.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR011096; Propep_M4_M36.
 DR InterPro; IPR005075; Propep_PepSY.
 DR Pfam; PF07504; FTP; 1.
 DR Pfam; PF03413; PepSY; 1.
 DR Pfam; PF01447; Peptidase_M4; 1.
 DR Pfam; PF02868; Peptidase_M4_C; 1.
 DR PRINTS; PR00730; THERMOLYSIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Signal; Virulence; Zinc; Zymogen.
 FT SIGNAL 1 24 Potential.
 FT PROPEP 25 200 Potential.
 FT CHAIN 201 510 Zinc metalloproteinase.
 FT METAL 349 349 Zinc (catalytic) (By similarity).
 FT ACT_SITE 350 350 By similarity.
 FT METAL 353 353 Zinc (catalytic) (By similarity).

FT METAL 373 373 Zinc (catalytic) (By similarity).
 FT ACT_SITE 437 437 Proton donor (By similarity).
 FT CONFLICT 47 47 T -> A (in Ref. 2).
 FT CONFLICT 103 103 T -> A (in Ref. 2).
 SQ SEQUENCE 510 AA; 57569 MW; C166CB56515BB175 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 510;
 Best Local Similarity 75.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8
 |||:| |
 Db 282 FYASEVYD 289

RESULT 14

Q6E9N9

ID Q6E9N9 PRELIMINARY; PRT; 510 AA.
 AC Q6E9N9;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Mpl.
 GN Name=mpl;
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL 33114, NRRL 33130, NRRL 33164, and NRRL 33090;
 RA Ward T.J., Gorski L., Borucki M.K., Mandrell R.E., Hutchins J.,
 RA Pupedis K.;
 RT "Intraspecific Phylogeny and Lineage Group Identification Based on the
 RT prfA Virulence Gene Cluster of Listeria monocytogenes.";
 RL J. Bacteriol. 186:4994-5002(2004).
 DR EMBL; AY512445; AAS85071.1; -.
 DR EMBL; AY512455; AAS85131.1; -.
 DR EMBL; AY512437; AAS85023.1; -.
 DR EMBL; AY512466; AAS85197.1; -.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001570; Peptidase_M4.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR011096; Propep_M4_M36.
 DR InterPro; IPR005075; Propep_PepSY.
 DR Pfam; PF07504; FTP; 1.
 DR Pfam; PF03413; PepSY; 1.
 DR Pfam; PF01447; Peptidase_M4; 1.
 DR Pfam; PF02868; Peptidase_M4_C; 1.
 DR PRINTS; PR00730; THERMOLYSIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 510 AA; 57581 MW; F8F1E03E89306A11 CRC64;

Query Match 71.1%; Score 32; DB 2; Length 510;
 Best Local Similarity 75.0%; Pred. No. 2.3e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8
|||:| |
Db 282 FYASEVYD 289

RESULT 15

Q6EA99

ID Q6EA99 PRELIMINARY; PRT; 510 AA.
AC Q6EA99;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Mpl.
GN Name=mpl;
OS *Listeria monocytogenes*.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 33032, NRRL 33033, NRRL 33038, NRRL 33068, NRRL 33073,
RC NRRL 33074, NRRL 33124, NRRL 33126, NRRL 33160, NRRL 33178,
RC NRRL 33186, NRRL 33218, and NRRL 33015;
RA Ward T.J., Gorski L., Borucki M.K., Mandrell R.E., Hutchins J.,
RA Papedis K.;
RT "Intraspecific Phylogeny and Lineage Group Identification Based on the
RT prfA Virulence Gene Cluster of *Listeria monocytogenes*.";
RL J. Bacteriol. 186:4994-5002(2004).
DR EMBL; AY512410; AAS84861.1; -.
DR EMBL; AY512411; AAS84867.1; -.
DR EMBL; AY512429; AAS84975.1; -.
DR EMBL; AY512431; AAS84987.1; -.
DR EMBL; AY512416; AAS84897.1; -.
DR EMBL; AY512432; AAS84993.1; -.
DR EMBL; AY512452; AAS85113.1; -.
DR EMBL; AY512473; AAS85239.1; -.
DR EMBL; AY512489; AAS85335.1; -.
DR EMBL; AY512403; AAS84821.1; -.
DR EMBL; AY512481; AAS85287.1; -.
DR EMBL; AY512465; AAS85191.1; -.
DR EMBL; AY512450; AAS85101.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001570; Peptidase_M4.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR011096; Propep_M4_M36.
DR InterPro; IPR005075; Propep_PepSY.
DR Pfam; PF07504; FTP; 1.
DR Pfam; PF03413; PepSY; 1.
DR Pfam; PF01447; Peptidase_M4; 1.
DR Pfam; PF02868; Peptidase_M4_C; 1.
DR PRINTS; PR00730; THERMOLYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 510 AA; 57539 MW; 662EA4CBB39E863A CRC64;

Query Match 71.1%; Score 32; DB 2; Length 510;
Best Local Similarity 75.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FYATEVXD 8
|||:| |
Db 282 FYASEVYD 289

Search completed: February 10, 2005, 15:57:35
Job time : 74.9577 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 113.338 Seconds
(without alignments)
44.362 Million cell updates/sec

Title: US-10-067-484-7
Perfect score: 65
Sequence: 1 MYATEVLDDLGSK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	65	100.0	13	5	ABB81974	Abb81974 30 kDa ra
2	39	60.0	275	2	AAR60476	Aar60476 Serine pr
3	39	60.0	721	6	ADA55497	Ada55497 Human pro
4	39	60.0	1308	6	ABR96139	Abr96139 Human NOV
5	39	60.0	1308	7	ADE28104	Ade28104 Human NTR
6	39	60.0	1308	7	ADI60123	Adi60123 Secreted
7	39	60.0	1309	6	ABR96140	Abr96140 Human NOV
8	38	58.5	145	7	ADC95858	Adc95858 E. faeciu
9	38	58.5	286	4	AAM25825	Aam25825 Human pro

10	38	58.5	286	4	AAM25950	Aam25950 Human pro
11	38	58.5	286	4	ABB11042	Abb11042 Human sec
12	38	58.5	436	8	ADS29240	Ads29240 Bacterial
13	38	58.5	436	8	ADS24427	Ads24427 Bacterial
14	38	58.5	546	5	ABB90404	Abb90404 Human pol
15	38	58.5	600	2	AAY29861	Aay29861 Human sec
16	38	58.5	658	7	ADE07881	Ade07881 Novel pro
17	38	58.5	704	2	AAY29517	Aay29517 Human lun
18	38	58.5	704	3	AAB44467	Aab44467 Human lun
19	38	58.5	704	4	AAE13809	Aae13809 Human lun
20	38	58.5	704	7	ADD66499	Add66499 Human lun
21	38	58.5	704	7	ADE87753	Ade87753 Human lun
22	38	58.5	1024	4	ABB65649	Abb65649 Drosophil
23	37	56.9	217	7	ADH87496	Adh87496 Enterococ
24	37	56.9	268	1	AAP90375	Aap90375 Subtilisi
25	37	56.9	268	4	ABM00020	Abm00020 Bacillus
26	37	56.9	272	6	ABR63750	Abr63750 Bacillus
27	37	56.9	272	6	ABR63734	Abr63734 Bacillus
28	37	56.9	272	7	ABR63719	Abr63719 Bacillus
29	37	56.9	272	7	ABR63645	Abr63645 Bacillus
30	37	56.9	280	8	ADI67211	Adi67211 Lactobaci
31	37	56.9	283	8	ADI67245	Adi67245 Lactobaci
32	37	56.9	316	2	AAY42795	Aay42795 Streptoco
33	37	56.9	316	5	ABP29264	Abp29264 Streptoco
34	37	56.9	361	2	AAR51233	Aar51233 Heat resi
35	37	56.9	374	2	AAR36727	Aar36727 Modified
36	37	56.9	374	2	AAW48410	Aaw48410 Bacillus
37	37	56.9	375	6	AAE29941	Aae29941 Bacillus
38	37	56.9	513	2	AAR04585	Aar04585 Aquaricin
39	37	56.9	513	2	AAR13181	Aar13181 T.aquatic
40	37	56.9	513	2	AAR67653	Aar67653 Aqualysin
41	37	56.9	948	6	ABU37952	Abu37952 Protein e
42	37	56.9	982	6	ABU37382	Abu37382 Protein e
43	37	56.9	1007	6	ABP80992	Abp80992 N. gonorr
44	36	55.4	13	5	ADG66421	Adg66421 B. amylo
45	36	55.4	15	7	ADB61329	Adb61329 P1 protea

ALIGNMENTS

RESULT 1

ABB81974

ID ABB81974 standard; peptide; 13 AA.

XX

AC ABB81974;

XX

DT 25-NOV-2002 (first entry)

XX

DE 30 kDa ragweed pollen allergen tryptic peptide 7.

XX

KW Ragweed; pollen; allergen; Ambt 7; glycoprotein; antiallergic;
KW immunotherapy; disulphide protein.

XX

OS Ambrosia elatior.

XX

PN WO200263012-A2.

XX
PD 15-AUG-2002.
XX
PF 04-FEB-2002; 2002WO-US003346.
XX
PR 05-FEB-2001; 2001US-0266686P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Buchanan BB, Del Val G, Frick OL;
XX
DR WPI; 2002-657539/70.
XX
PT New ragweed pollen allergens, useful in allergy testing and immunotherapy
PT regimens, particularly for treating sensitivity to pollen or pollen
PT allergy (e.g. anaphylaxis, or symptoms of hives or asthma) in a mammal,
PT especially a human.
XX
PS Claim 1; Page 53; 70pp; English.
XX
CC The invention relates to an isolated pollen allergen purified from
CC ragweed pollen, substantially free of any other pollen proteins, or a
CC protein that is an antigenic fragment of a pollen allergen Ambt 7. The
CC allergen is characterized by the following physiochemical and biological
CC properties: (a) being contained in pollen extracts; (b) a glycoprotein;
CC (c) a sulphhydryl group containing protein; (d) a molecular weight of
CC about 30 kDa as determined by SDS-polyacrylamide gel electrophoresis; and
CC (e) possessing allergen activity. The pollen allergen, or antigenic
CC protein fragment of the pollen allergen Ambt 7, or composition is useful
CC for treating sensitivity to pollen or pollen allergy in a mammal. This
CC allergy includes anaphylaxis or atopy, which includes the symptoms of hay
CC fever, asthma or hives. The allergen is also useful in allergy testing
CC and immunotherapy regimens. Sequences ABB81968-978 represent tryptic
CC peptide fragments of the 30 kDa ragweed complete pollen extract
CC disulphide protein allergen
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 65; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYATEVLDLDGSK 13
|||||||
Db 1 MYATEVLDLDGSK 13

RESULT 2

AAR60476

ID AAR60476 standard; protein; 275 AA.

XX

AC AAR60476;

XX

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 03-APR-1995 (first entry)

XX

DE Serine protease of *Bacillus subtilis* 168.
 XX
 KW Serine protease; protease; enzyme; peptide ligase; variant;
 KW *Bacillus subtilis*; modification.
 XX
 OS *Bacillus* sp.
 XX
 PN WO9418329-A2.
 XX
 PD 18-AUG-1994.
 XX
 PF 02-FEB-1994; 94WO-US001336.
 XX
 PR 04-FEB-1993; 93US-00013445.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Abrahmsen L, Burnier J, Wells JA, Jackson DY;
 XX
 DR WPI; 1994-279750/34.
 XX
 PT New serine protease variants - having amino acid modifications to improve
 PT peptide ligase activity in the synthesis of polypeptide(s).
 XX
 PS Disclosure; Fig 6; 62pp; English.
 XX
 CC Serine protease variants with greater peptide ligase activity than wild
 CC type counterparts may be produced by changing at least two amino acids in
 CC a serine protease precursor. The changes comprise (1) the replacement or
 CC modification of a side chain of an active site serine residue in the
 CC precursor protease to substitute the nucleophilic oxygen of the side
 CC chain with a different nucleophile and (2) the replacement or
 CC modification of the side chain of a second amino acid residue in the
 CC precursor protease. (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 275 AA;

Query Match 60.0%; Score 39; DB 2; Length 275;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MYATEVLDLGGS 12
 :|| |||| ||
 Db 90 LYAVEVL DSTGS 101

RESULT 3
 ADA55497

ID ADA55497 standard; protein; 721 AA.
 XX
 AC ADA55497;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human protein, SEQ ID 3065.
 XX

KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW inflammatory disease; osteoporosis; neurological disease.
 XX
 OS Homo sapiens.
 XX
 PN EP1293569-A2.
 XX
 PD 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-00006586.
 XX
 PR 14-SEP-2001; 2001JP-00328381.
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai-K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 DR WPI; 2003-395539/38.
 DR N-PSDE; ADA53858.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 PS Claim 14; SEQ ID NO 3065; 205pp; English.
 XX
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX
 SQ Sequence 721 AA;

Query Match 60.0%; Score 39; DB 6; Length 721;
 Best Local Similarity 70.0%; Pred. No. 2.9e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 YATEVLDLDG 11
 | :||:||||
 Db 154 YRSEVVDLDG 163

RESULT 4
 ABR96139
 ID ABR96139 standard; protein; 1308 AA.
 XX
 AC ABR96139;
 XX
 DT 15-SEP-2003 (first entry)
 XX
 DE Human NOV7a protein SEQ ID NO:20.

XX
KW Human; NOVX; G protein-coupled receptor; cytostatic; cardiovascular;
KW immunosuppressive; anti-HIV; antiasthmatic; antiarteriosclerotic; AIDS;
KW hypotensive; gene therapy; cardiomyopathy; atherosclerosis; hypertension;
KW congenital heart defect; aortic stenosis; atrial septal defect; neoplasm;
KW atrioventricular canal defect; pulmonary stenosis; prostate cancer;
KW uterine cancer; graft versus host disease; multiple sclerosis; GPCR;
KW acquired immunodeficiency syndrome; Crohn's disease; bronchial asthma;
KW chromosome mapping; forensic identification.

XX
OS Homo sapiens.

XX
PN WO200290568-A2.

XX
PD 14-NOV-2002.

XX
PF 02-MAY-2002; 2002WO-US014341.

XX
PR 03-MAY-2001; 2001US-0288935P.
PR 07-MAY-2001; 2001US-0289087P.
PR 08-MAY-2001; 2001US-0289620P.
PR 08-MAY-2001; 2001US-0289621P.
PR 09-MAY-2001; 2001US-0289817P.
PR 09-MAY-2001; 2001US-0289818P.
PR 11-MAY-2001; 2001US-0290194P.
PR 14-MAY-2001; 2001US-0290753P.
PR 15-MAY-2001; 2001US-0291189P.
PR 16-MAY-2001; 2001US-0291243P.
PR 18-MAY-2001; 2001US-0292001P.
PR 21-MAY-2001; 2001US-0292374P.
PR 22-MAY-2001; 2001US-0292587P.
PR 23-MAY-2001; 2001US-0293107P.
PR 24-MAY-2001; 2001US-0293589P.
PR 25-MAY-2001; 2001US-0293747P.
PR 29-MAY-2001; 2001US-0294110P.
PR 30-MAY-2001; 2001US-0294434P.
PR 14-AUG-2001; 2001US-0312192P.
PR 17-AUG-2001; 2001US-0313173P.
PR 17-AUG-2001; 2001US-0313187P.
PR 12-SEP-2001; 2001US-0318728P.
PR 12-SEP-2001; 2001US-0318744P.
PR 15-NOV-2001; 2001US-0335910P.
PR 28-NOV-2001; 2001US-0333891P.
PR 28-NOV-2001; 2001US-0333942P.
PR 03-JAN-2002; 2002US-0345776P.
PR 04-JAN-2002; 2002US-0345220P.
PR 01-MAY-2002; 2002US-00136071.

XX
PA (CURA-) CURAGEN CORP.

XX
PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
PI Edinger SR, Ellerman K, Gangolli EA, Gerlach VL, Gorman L;
PI Gunther E, Herrmann JL, Ji W, Lepley DM, Lewin DA, Li L;
PI Macdougall JR, Malyankar UM, Mezes PD, Padigar M, Patturajan M;
PI Peyman JA, Rastelli L, Rieger DK, Rothenberg ME, Shenoy SG;
PI Smithson G, Spytek KA, Stone DJ, Taupier RJ, Tchernev VT;
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong H, Miller CE;

XX
DR WPI; 2003-111987/10.
DR N-PSDB; ACF16948.
XX
PT New NOVX polypeptides and polynucleotides useful for treating or
PT preventing e.g. cardiomyopathy, atherosclerosis, hypertension, congenital
PT heart defects, aortic stenosis, atrial septal defect, or atrioventricular
PT canal defect.
XX
PS Claim 1; Page 118; 491pp; English.
XX
CC ACF16939 to ACF17000 encode the human G protein-coupled receptor (GPCR)
CC proteins, designated NOVX proteins, given in ABR96130 to ABR96191. The
CC NOVX sequences can have cytostatic, cardiovascular, antiasthmatic,
CC immunosuppressive, anti-HIV (human immunodeficiency virus), hypotensive
CC and antiarteriosclerotic activities, and can be used in gene therapy.
CC NOVX polypeptides can be used for treating a syndrome associated with a
CC human disease such as a pathology associated with the polypeptide. NOVX
CC polypeptides, polynucleotides and antibodies can be used for treating or
CC preventing e.g. cardiomyopathy, atherosclerosis, hypertension, congenital
CC heart defects, aortic stenosis, atrial septal defect, atrioventricular
CC canal defect, pulmonary stenosis, prostate cancer, uterine cancer,
CC neoplasm, graft versus host disease, acquired immunodeficiency syndrome
CC (AIDS), Crohn's disease, multiple sclerosis, or bronchial asthma. The
CC nucleic acid sequences may be used in chromosome mapping, identifying
CC individual from minute biological samples (tissue typing), and in
CC forensic identification of a biological sample. ACF17001 to ACF17117
CC represent PCR primers and probes for the NOVX sequences, which are used
CC in an example from the present invention
XX
SQ Sequence 1308 AA;

Query Match 50.0%; Score 39; DB 6; Length 1308;
Best Local Similarity 70.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YATEVLDLDG 11
| :||:||||
Db 179 YRSEVVDLDG 188

RESULT 5

ADE28104

ID ADE28104 standard; protein; 1308 AA.

XX

AC ADE28104;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human NTRAN protein - SEQ ID 9.

XX

KW human; neurotransmission-associated protein; NTRAN; cytostatic;
KW immunomodulator; immune disorder; cancer; gene therapy.

XX

OS Homo sapiens.

XX

PN WO2003051902-A1.

XX
 PD 26-JUN-2003.
 XX
 PF 12-DEC-2002; 2002WO-US040059.
 XX
 PR 14-DEC-2001; 2001US-0340798P.
 PR 18-MAR-2002; 2002US-0365645P.
 PR 25-MAR-2002; 2002US-0367662P.
 PR 10-MAY-2002; 2002US-0379887P.
 PR 31-MAY-2002; 2002US-0384639P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Baughn MR, Bhatia U, Blake JJ, Burrill JD, Elliott VS;
 PI Emerling BM, Forsythe IJ, Gietzen KJ, Görvad AE, Griffin JA;
 PI Hafalia AJA, Ho A, Jackson AA, Jiang X, Kable AE, Kearney L;
 PI Khare R, Lee EA, Lee S, Lu DAM, Marquis JP, Lehr-Mason PM;
 PI Ramkumar J, Richardson TW, Sprague WW, Tran UK, Chawla NK;
 PI Warren BA, Yue H, Zheng W;
 XX
 DR WPI; 2003-514037/48.
 DR N-PSDB; ADE28126.
 XX
 PT New human neurotransmission-associated proteins (NTRAN) polypeptide,
 PT useful for preparing a composition for treating a disease associated with
 PT decreased expression or overexpression of NTRAN e.g., cancer.
 XX
 PS Claim 1; SEQ ID NO 9; 261pp; English.
 XX
 CC The invention relates to a novel isolated human neurotransmission-
 CC associated proteins (NTRAN) polypeptide. The polypeptide of the invention
 CC demonstrates cytostatic and immunomodulator activities and may be useful
 CC for preparing a composition for diagnosing or treating a disease or
 CC condition associated with decreased expression or overexpression of
 CC functional NTRAN including immune disorders or cancer, as well as during
 CC gene therapy procedures. The current sequence is that of the human NTRAN
 CC protein of the invention.
 XX
 SQ Sequence 1308 AA;

Query Match 60.0%; Score 39; DB 7; Length 1308;
 Best Local Similarity 70.0%; Pred. No. 5.5e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YATEVLDLDG 11
 | :||:||||
 Db 179 YRSEVVDLDG 188

RESULT 6
 ADI60123
 ID ADI60123 standard; protein; 1308 AA.
 XX
 AC ADI60123;
 XX
 DT 15-APR-2004 (first entry)
 XX

DE Secreted polypeptide #7.
 XX
 KW osteopathic; vulnerary; cytostatic; gene therapy; diagnosis; forensics;
 KW gene mapping; mutation identification; biodiversity; chromosome marker;
 KW immune response; myeloid cell disorder; lymphoid cell disorder;
 KW bone cartilage; tendon; ligament; nerve tissue growth; wound healing;
 KW burns; incision; ulcer; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO2003025142-A2.
 XX
 PD 27-MAR-2003.
 XX
 PF 18-SEP-2002; 2002WO-US029636.
 XX
 PR 18-SEP-2001; 2001US-0323349P.
 PR 16-SEP-2002; 2002US-00323349.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT;
 XX
 DR WPI; 2003-354601/33.
 DR N-PSDB; ADI60468.
 XX
 PT New polynucleotides and secreted proteins, useful for treating myeloid or
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
 PT tissue growth or regeneration, in wound healing, and in tissue repair and
 PT replacement.
 XX
 PS Claim 20; SEQ ID NO 158; 243pp; English.
 XX
 CC The invention relates to novel isolated polynucleotides or a sequence
 CC encoding a polypeptide with biological activity, where the polynucleotide
 CC hybridizes to the polynucleotide under stringent hybridization conditions
 CC or has greater than 99% sequence identity with the polynucleotide. The
 CC polynucleotides and polypeptides are useful in diagnostics, forensics,
 CC gene mapping, identification of mutations responsible for genetic
 CC disorders and other traits, to assess biodiversity, as nutritional
 CC sources or supplements. The polynucleotides may also be used as molecular
 CC weight markers, chromosome markers or map related gene positions, or as
 CC an antigen to raise anti-DNA antibodies or elicit immune response. The
 CC polypeptides are useful for raising antibodies, as markers for tissues in
 CC which the corresponding polypeptide is expressed, for re-engineering
 CC damaged or diseased tissues, for treating myeloid or lymphoid cell
 CC disorders, in bone cartilage, tendon, ligament and/or nerve tissue growth
 CC or regeneration, in wound healing, in tissue repair and replacement, in
 CC healing of burns, incisions and ulcers, and in treating cancer. This
 CC sequence corresponds to a protein sequence of the invention.
 XX
 SQ Sequence 1308 AA;

Query Match 60.0%; Score 39; DB 7; Length 1308;
 Best Local Similarity 70.0%; Pred. No. 5.5e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YATEVLDDLDG 11
| :||:||||
Db 179 YRSEVVDLDG 188

RESULT 7

ABR96140

ID ABR96140 standard; protein; 1309 AA.

XX

AC ABR96140;

XX

DT 15-SEP-2003 (first entry)

XX

DE Human NOV7b protein SEQ ID NO:22.

XX

KW Human; NOVX; G protein-coupled receptor; cytostatic; cardiovascular;
KW immunosuppressive; anti-HIV; antiasthmatic; antiarteriosclerotic; AIDS;
KW hypotensive; gene therapy; cardiomyopathy; atherosclerosis; hypertension;
KW congenital heart defect; aortic stenosis; atrial septal defect; neoplasm;
KW atrioventricular canal defect; pulmonary stenosis; prostate cancer;
KW uterine cancer; graft versus host disease; multiple sclerosis; GPCR;
KW acquired immunodeficiency syndrome; Crohn's disease; bronchial asthma;
KW chromosome mapping; forensic identification.

XX

OS Homo sapiens.

XX

PN WO200290568-A2.

XX

PD 14-NOV-2002.

XX

PF 02-MAY-2002; 2002WO-US014341.

XX

PR 03-MAY-2001; 2001US-0288935P.

PR 07-MAY-2001; 2001US-0289087P.

PR 08-MAY-2001; 2001US-0289620P.

PR 08-MAY-2001; 2001US-0289621P.

PR 09-MAY-2001; 2001US-0289817P.

PR 09-MAY-2001; 2001US-0289818P.

PR 11-MAY-2001; 2001US-0290194P.

PR 14-MAY-2001; 2001US-0290753P.

PR 15-MAY-2001; 2001US-0291189P.

PR 16-MAY-2001; 2001US-0291243P.

PR 18-MAY-2001; 2001US-0292001P.

PR 21-MAY-2001; 2001US-0292374P.

PR 22-MAY-2001; 2001US-0292587P.

PR 23-MAY-2001; 2001US-0293107P.

PR 24-MAY-2001; 2001US-0293589P.

PR 25-MAY-2001; 2001US-0293747P.

PR 29-MAY-2001; 2001US-0294110P.

PR 30-MAY-2001; 2001US-0294434P.

PR 14-AUG-2001; 2001US-0312192P.

PR 17-AUG-2001; 2001US-0313173P.

PR 17-AUG-2001; 2001US-0313187P.

PR 12-SEP-2001; 2001US-0318728P.

PR 12-SEP-2001; 2001US-0318744P.

PR 15-NOV-2001; 2001US-0335910P.

PR 28-NOV-2001; 2001US-0333891P.
PR 28-NOV-2001; 2001US-0333942P.
PR 03-JAN-2002; 2002US-0345776P.
PR 04-JAN-2002; 2002US-0345220P.
PR 01-MAY-2002; 2002US-00136071.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
PI Edinger SR, Ellerman K, Gangolli EA, Gerlach VL, Gorman L;
PI Gunther E, Herrmann JL, Ji W, Lepley DM, Lewin DA, Li L;
PI Macdougall JR, Malyankar UM, Mezes PD, Padigar M, Patturajan M;
PI Peyman JA, Rastelli L, Rieger DK, Rothenberg ME, Shenoy SG;
PI Smithson G, Spytek KA, Stone DJ, Taupier RJ, Tchernev VT;
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong H, Miller CE;

XX

DR WPI; 2003-111987/10.

DR N-PSDB; ACF16949.

XX

PT New NOVX polypeptides and polynucleotides useful for treating or
PT preventing e.g. cardiomyopathy, atherosclerosis, hypertension, congenital
PT heart defects, aortic stenosis, atrial septal defect, or atrioventricular
PT canal defect.

XX

PS Claim 1; Page 120; 491pp; English.

XX

CC ACF16939 to ACF17000 encode the human G protein-coupled receptor (GPCR)
CC proteins, designated NOVX proteins, given in ABR96130 to ABR96191. The
CC NOVX sequences can have cytostatic, cardiovascular, antiasthmatic,
CC immunosuppressive, anti-HIV (human immunodeficiency virus), hypotensive
CC and antiarteriosclerotic activities, and can be used in gene therapy.
CC NOVX polypeptides can be used for treating a syndrome associated with a
CC human disease such as a pathology associated with the polypeptide. NOVX
CC polypeptides, polynucleotides and antibodies can be used for treating or
CC preventing e.g. cardiomyopathy, atherosclerosis, hypertension, congenital
CC heart defects, aortic stenosis, atrial septal defect, atrioventricular
CC canal defect, pulmonary stenosis, prostate cancer, uterine cancer,
CC neoplasm, graft versus host disease, acquired immunodeficiency syndrome
CC (AIDS), Crohn's disease, multiple sclerosis, or bronchial asthma. The
CC nucleic acid sequences may be used in chromosome mapping, identifying
CC individual from minute biological samples (tissue typing), and in
CC forensic identification of a biological sample. ACF17001 to ACF17117
CC represent PCR primers and probes for the NOVX sequences, which are used
CC in an example from the present invention

XX

SQ Sequence 1309 AA;

Query Match 60.0%; Score 39; DB 6; Length 1309;
Best Local Similarity 70.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YATEVLDLDG 11
| :||:||||
Db 179 YRSEVVDLDG 188

RESULT 8

ADC95858

ID ADC95858 standard; protein; 145 AA.

XX

AC ADC95858;

XX

DT 01-JAN-2004 (first entry)

XX

DE E. faecium protein sequence SEQ ID 5485.

XX

KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.

XX

OS Enterococcus faecium.

XX

PN US6583275-B1.

XX

PD 24-JUN-2003.

XX

PF 30-JUN-1998; 98US-00107532.

XX

PR 02-JUL-1997; 97US-0051571P.

PR 14-MAY-1998; 98US-0085598P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Doucette-Stamm LA, Bush D;

XX

DR WPI; 2003-799836/75.

DR N-PSDB; ADC92204.

XX

PT New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.

XX

PS Example 1; SEQ ID NO 5485; 243pp; English.

XX

CC The invention relates to an isolated nucleic acid derived from
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridising to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC transcription regulatory element, a cell comprising the vector and a
CC single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids is useful for diagnosing pathological conditions
CC resulting from E. faecium bacterial infection (e.g. urinary tract
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of Candida albicans -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating Enterococcus faecium infections. The present sequence represents
CC one if the disclosed E. faecium proteins.

XX

SQ Sequence 145 AA;

Query Match 58.5%; Score 38; DB 7; Length 145;
Best Local Similarity 58.3%; Pred. No. 78;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MYATEVLDDLGS 12
|:: || |||:
Db 125 MFSLEVQDLGN 136

RESULT 9

AAM25825

ID AAM25825 standard; protein; 286 AA.

XX

AC AAM25825;

XX

DT 16-OCT-2001 (first entry)

XX

DE Human protein sequence SEQ ID NO:1340.

XX

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;

KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;

KW antibacterial; endocrine; cardiant; central nervous system; virucide;

KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;

KW antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;

KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;

KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;

KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;

KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;

KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;

KW genetic disease; haematopoietic disorder; platelet disorder; asthma;

KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;

KW allergic rhinitis; diabetes; multiple sclerosis; depression;

KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;

KW neurological disorder.

XX

OS Homo sapiens.

XX

PN WO200153455-A2.

XX

PD 26-JUL-2001.

XX

PF 22-DEC-2000; 2000WO-US035017.

XX

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-457603/49.

DR N-PSDB; AAH99766.

XX

PT Isolated human polynucleotides encoding polypeptides, useful for the
treatment and diagnosis of e.g. cancer, ulcers and HIV infection.

XX
PS Claim 20; Page 278; 1217pp; English.
XX
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production, The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX
SQ Sequence 286 AA;

Query Match 58.5%; Score 38; DB 4; Length 286;
Best Local Similarity 53.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MYATEVLDLGSK 13
:| :| ||| |
Db 202 LYLKDVQDLGGK 214

RESULT 10

AAM25950

ID AAM25950 standard; protein; 286 AA.

XX

AC AAM25950;

XX

DT 16-OCT-2001 (first entry)

XX

DE Human protein sequence SEQ ID NO:1465.

XX

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;

KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000WO-US035017.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-457603/49.
 DR N-PSDB; AAH99891.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
 XX
 PS Claim 20; Page 293; 1217pp; English.
 XX
 CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
 CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production, The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders
 XX
 SQ Sequence 286 AA;

Query Match 58.5%; Score 38; DB 4; Length 286;
 Best Local Similarity 53.8%; Pred. No. 1.6e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGGSK 13
:| :| |||||
Db 202 LYLKDVQDLGGK 214

RESULT 11

ABB11042

ID ABB11042 standard; peptide; 286 AA.

XX

AC ABB11042;

XX

DT 11-JAN-2002 (first entry)

XX

DE Human secreted protein homologue, SEQ ID NO:1412.

XX

KW Human; cytokine; cell proliferation; cell differentiation; growth factor;

KW haematopoiesis regulation; tissue growth; immunomodulator; activin;

KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;

KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;

KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;

KW chronic inflammatory condition; proliferative retinopathy;

KW atherosclerosis; coronary heart disease; arterial ischaemia;

KW bone disorder; osteoporosis; vascular growth disorder;

KW tissue regeneration; wound healing; infection; immune disorder;

KW cell culture; drug screening; gene therapy; antiinflammatory;

KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;

KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;

KW antifungal; vulnerary; antiulcer.

XX

OS Homo sapiens.

XX

PN WO200157188-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US003800.

XX

PR 03-FEB-2000; 2000US-00496914.

PR 27-APR-2000; 2000US-00560875.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-457740/49.

DR N-PSDB; ABA08286.

XX

PT Human proteins and DNA encoding sequences useful for preventing, treating

PT or ameliorating a medical condition in a mammalian subject e.g. arthritis

PT and cancer.

XX

PS Claim 20; Page 139; 1963pp; English.

XX

CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and

CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The

CC invention also relates to vectors and recombinant host cells comprising a

CC nucleotide of the invention, methods of producing the novel polypeptides,

CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention
 XX
 SQ Sequence 286 AA;

Query Match 58.5%; Score 38; DB 4; Length 286;
 Best Local Similarity 53.8%; Pred. No. 1.6e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MYATEVLDLGSK 13
 :| :| |||||
 Db 202 LYLKDVQDLGGK 214

RESULT 12

ADS29240

ID ADS29240 standard; protein; 436 AA.

XX

AC ADS29240;

XX

DT 02-DEC-2004 (first entry)

XX

DE Bacterial polypeptide #18273.

XX

KW Recombinant DNA construct; transformed plant; improved plant property;

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 18273; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX

SQ Sequence 436 AA;

Query Match 58.5%; Score 38; DB 8; Length 436;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MYATEVLDDL 10
|| |||||
Db 146 MYNREVLDDL 155

RESULT 13

ADS24427

ID ADS24427 standard; protein; 436 AA.

XX

AC ADS24427;

XX

DT 02-DEC-2004 (first entry)

XX

DE Bacterial polypeptide #13460.

XX

KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX

OS Bacteria.

XX

PN US2003233675-A1.

XX

PD 18-DEC-2003.

XX

PF 20-FEB-2003; 2003US-00369493.

XX

PR 21-FEB-2002; 2002US-0360039P.

XX

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX

DR WPI; 2004-061375/06.

XX

PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

XX

PS Claim 1; SEQ ID NO 13460; 122pp; English.

XX

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX

SQ Sequence 436 AA;

Query Match 58.5%; Score 38; DB 8; Length 436;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDDL 10
|| |||||
Db 146 MYNREVLDDL 155

RESULT 14

ABB90404

ID ABB90404 standard; protein; 546 AA.

XX

AC ABB90404;

XX

DT 24-MAY-2002 (first entry)

XX

DE Human polypeptide SEQ ID NO 2780.

XX

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.

XX

OS Homo sapiens.

XX

PN WO200190304-A2.

XX

PD 29-NOV-2001.

XX

PF 18-MAY-2001; 2001WO-US016450.

XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR N-PSDB; ABL90813.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 11; SEQ ID NO 2780; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 546 AA;

Query Match 58.5%; Score 38; DB 5; Length 546;
Best Local Similarity 53.8%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MYATEVLDDGSK 13
:| :| |||||
Db 133 LYLKDVQDLDDGK 145

RESULT 15

AAAY29861

ID AAY29861 standard; protein; 600 AA.

XX

AC AAY29861;

XX

DT 17-NOV-1999 (first entry)

XX

DE Human secreted protein clone cb98_4.

XX

KW Human; secreted protein; biological activity; nutritional; cytokine;

KW cell proliferation; differentiation; immune stimulating; vaccine;
KW haematopoiesis regulation; tissue growth; haemostatic; thrombolytic;
KW anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 99
FT /note= "unspecified"
XX
PN WO9946287-A1.
XX
PD 16-SEP-1999.
XX
PF 11-MAR-1999; 99WO-US005243.
XX
PR 11-MAR-1998; 98US-0077521P.
PR 14-MAY-1998; 98US-00079124.
PR 10-MAR-1999; 99US-00266105.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Agostino MJ, Steininger RJ;
XX
DR WPI; 1999-551362/46.
DR N-PSDB; AAZ21093.
XX
PT Polynucleotides encoding secreted human proteins, derived from human
PT fetal brain, human adult blood, human adult bladder, or human adult
PT neural tissue cDNA libraries.
XX
PS Claim 9; Page 99-101; 118pp; English.
XX
CC AAZ21093 to AAZ21102 encode new human secreted proteins and AAY29861 to
CC AAY29873 represent the secreted proteins encoded by the polynucleotide
CC sequences. AAZ21103 to AAZ21112 represent probes for the secreted
CC proteins. The polynucleotides and proteins are predicted to have
CC biological activities which would make them suitable for treating,
CC preventing or ameliorating medical conditions in humans and animals,
CC although no supporting data is given. Suggested activities include
CC nutritional activity, cytokine and cell proliferation/differentiation
CC activity, immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides and proteins can also be used as
CC nutritional sources or supplements. Such uses include use as a protein or
CC amino acid supplement, use as a carbon source, use as a nitrogen source
CC and use as a source of carbohydrate. They may also have utility in
CC compositions used for bone, cartilage, tendon, ligament, and/or nerve
CC tissue growth or regeneration, as well as for wound healing and tissue
CC repair and replacement, and in the treatment of burns, incisions and
CC ulcers. The proteins which induce cartilage and/or bone growth in
CC circumstances where bone is not normally formed, have application in the
CC healing of bone fractures and cartilage damage or defects in humans and

CC other animals
XX
SQ Sequence 600 AA;

Query Match 58.5%; Score 38; DB 2; Length 600;
Best Local Similarity 53.8%; Pred. No. 3.6e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGGSK 13
:| :| ||| |
Db 176 LYLKDVQDLGGK 188

Search completed: February 10, 2005, 15:48:44
Job time : 115.338 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 29.1127 Seconds
(without alignments)
33.334 Million cell updates/sec

Title: US-10-067-484-7
Perfect score: 65
Sequence: 1 MYATEVLDDLGGSK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	39	60.0	161	4	US-09-902-540-10892	Sequence 10892, A	
2	38	58.5	145	4	US-09-107-532A-5485	Sequence 5485, Ap	
3	38	58.5	704	4	US-09-370-838-191	Sequence 191, App	
4	38	58.5	704	4	US-09-854-133-191	Sequence 191, App	
5	37	56.9	162	4	US-09-765-815-14	Sequence 14, Appl	
6	37	56.9	217	4	US-09-134-000C-5381	Sequence 5381, Ap	
7	37	56.9	268	1	US-08-431-387-4	Sequence 4, Appli	
8	37	56.9	268	4	US-10-310-730-2	Sequence 2, Appli	
9	37	56.9	280	4	US-09-634-238-303	Sequence 303, App	
10	36	55.4	245	4	US-09-489-039A-11571	Sequence 11571, A	
11	36	55.4	247	4	US-09-489-039A-8478	Sequence 8478, Ap	
12	36	55.4	268	4	US-09-512-251A-2	Sequence 2, Appli	
13	36	55.4	268	4	US-09-515-150A-2	Sequence 2, Appli	
14	36	55.4	268	4	US-09-196-281-5	Sequence 5, Appli	
15	36	55.4	268	4	US-10-007-389-2	Sequence 2, Appli	
16	36	55.4	269	1	US-08-566-369-10	Sequence 10, Appl	
17	36	55.4	269	1	US-08-566-369-13	Sequence 13, Appl	
18	36	55.4	269	3	US-09-074-331-10	Sequence 10, Appl	
19	36	55.4	269	3	US-09-074-331-13	Sequence 13, Appl	
20	36	55.4	269	5	PCT-US95-01937-10	Sequence 10, Appl	
21	36	55.4	269	5	PCT-US95-01937-13	Sequence 13, Appl	
22	36	55.4	273	4	US-09-088-912-1	Sequence 1, Appli	
23	36	55.4	275	1	US-08-431-387-3	Sequence 3, Appli	
24	36	55.4	275	1	US-08-322-677A-7	Sequence 7, Appli	
25	36	55.4	275	1	US-08-322-676-7	Sequence 7, Appli	
26	36	55.4	275	1	US-08-460-343B-72	Sequence 72, Appl	
27	36	55.4	275	1	US-08-460-343B-74	Sequence 74, Appl	
28	36	55.4	275	1	US-08-398-028B-72	Sequence 72, Appl	
29	36	55.4	275	1	US-08-398-028B-74	Sequence 74, Appl	
30	36	55.4	275	2	US-08-504-265B-72	Sequence 72, Appl	
31	36	55.4	275	2	US-08-504-265B-90	Sequence 90, Appl	
32	36	55.4	275	2	US-08-140-083A-9	Sequence 9, Appli	
33	36	55.4	275	2	US-08-865-203-8	Sequence 8, Appli	
34	36	55.4	275	2	US-09-135-658-3	Sequence 3, Appli	
35	36	55.4	275	2	US-07-849-420-8	Sequence 8, Appli	
36	36	55.4	275	3	US-08-898-218-7	Sequence 7, Appli	
37	36	55.4	275	3	US-08-848-793-7	Sequence 7, Appli	
38	36	55.4	275	3	US-09-253-854-8	Sequence 8, Appli	
39	36	55.4	275	3	US-08-955-424-8	Sequence 8, Appli	
40	36	55.4	275	3	US-09-178-155-3	Sequence 3, Appli	
41	36	55.4	275	3	US-09-445-270-2	Sequence 2, Appli	
42	36	55.4	275	3	US-09-467-536A-2	Sequence 2, Appli	
43	36	55.4	275	3	US-09-234-957-2	Sequence 2, Appli	
44	36	55.4	275	4	US-08-394-011-1	Sequence 1, Appli	
45	36	55.4	275	4	US-08-397-329-1	Sequence 1, Appli	

ALIGNMENTS

RESULT 1

US-09-902-540-10892

; Sequence 10892, Application US/09902540

```

; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10892
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10892

```

```

Query Match          60.0%; Score 39; DB 4; Length 161;
Best Local Similarity 53.8%; Pred. No. 10;
Matches      7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 MYATEVLDDL D GSK 13
        :|||  |||:|::
Db      97 LYATGFLDLEGTE 109

```

RESULT 2

US-09-107-532A-5485

```

; Sequence 5485, Application US/09107532A
; Patent No. 6583275

```

GENERAL INFORMATION:

```

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND
THERAPEUTICS

```

```

; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571

```

```

; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5485:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...145
; SEQUENCE DESCRIPTION: SEQ ID NO: 5485:

```

US-09-107-532A-5485

```

Query Match          58.5%;  Score 38;  DB 4;  Length 145;
Best Local Similarity 58.3%;  Pred. No. 14;
Matches      7;  Conservative  3;  Mismatches    2;  Indels    0;  Gaps    0;

```

```

QY      1 MYATEVLDDLGS 12
        |:: || ||||:
Db      125 MFSLEVQDLGN 136

```

RESULT 3

```

US-09-370-838-191
; Sequence 191, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 191
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-370-838-191

```

```

Query Match          58.5%;  Score 38;  DB 4;  Length 704;

```

Best Local Similarity 53.8%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDGSK 13
:| :| |||||
Db 280 LYLKDVQDLGGK 292

RESULT 4

US-09-854-133-191

; Sequence 191, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 191
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-191

Query Match 58.5%; Score 38; DB 4; Length 704;
Best Local Similarity 53.8%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDGSK 13
:| :| |||||
Db 280 LYLKDVQDLGGK 292

RESULT 5

US-09-765-815-14

; Sequence 14, Application US/09765815
; Patent No. 6673586
; GENERAL INFORMATION:
; APPLICANT: Balk, Steven
; TITLE OF INVENTION: No. 6673586el Steroid Hormone Receptor
; TITLE OF INVENTION: Interacting Protein Kinase
; FILE REFERENCE: 01948/068002
; CURRENT APPLICATION NUMBER: US/09/765,815
; CURRENT FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/176,859
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 162

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-765-815-14

Query Match 56.9%; Score 37; DB 4; Length 162;
Best Local Similarity 58.3%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGS 12
:|||||:|
Db 103 LYATEVVDFS 114

RESULT 6

US-09-134-000C-5381
; Sequence 5381, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5381
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5381

Query Match 56.9%; Score 37; DB 4; Length 217;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGS 12
|||:||||:
Db 26 MYQTILFDLDGT 37

RESULT 7

US-08-431-387-4
; Sequence 4, Application US/08431387
; Patent No. 5677163
; GENERAL INFORMATION:
; APPLICANT: Mainzer, Stanley E.
; APPLICANT: Lad, Pushkaraj J.
; APPLICANT: Schmidt, Brian
; TITLE OF INVENTION: Cleaning Compositions Containing
; TITLE OF INVENTION: No. 5677163el Alkaline Proteases
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 180 Kimball Way

; CITY: South San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/431,387
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/950,856A
 ; FILING DATE: September 24, 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Horn, Margaret A.
 ; REGISTRATION NUMBER: 33,401
 ; REFERENCE/DOCKET NUMBER: GC224
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 742-7536
 ; TELEFAX: (415) 742-7217
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 268 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-431-387-4

Query Match 56.9%; Score 37; DB 1; Length 268;
 Best Local Similarity 58.3%; Pred. No. 44;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MYATEVLDDLGS 12
 :|| :||| :||
 Db 87 LYAVKVLDRNGS 98

RESULT 8

US-10-310-730-2

; Sequence 2, Application US/10310730
 ; Patent No. 6835821
 ; GENERAL INFORMATION:
 ; APPLICANT: Hastrup, Sven
 ; APPLICANT: Branner, Sven
 ; APPLICANT: Horris, Fanny
 ; APPLICANT: Petersen, Steffen
 ; APPLICANT: No. 6835821skov-Lauridsen, Leif
 ; APPLICANT: Jensen, Villy
 ; APPLICANT: Aaslyng, Dorrit
 ; TITLE OF INVENTION: Useful Mutations of Bacterial Alkaline Protease
 ; FILE REFERENCE: 3160.250-US
 ; CURRENT APPLICATION NUMBER: US/10/310,730
 ; CURRENT FILING DATE: 2002-12-05

; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Bacillus
US-10-310-730-2

Query Match 56.9%; Score 37; DB 4; Length 268;
Best Local Similarity 58.3%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDLGDS 12
: || : ||| : ||
Db 87 LYAVKVLDRNGS 98

RESULT 9

US-09-634-238-303

; Sequence 303, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-303

Query Match 56.9%; Score 37; DB 4; Length 280;
Best Local Similarity 53.8%; Pred. No. 46;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDLGSK 13
| | : ||| : || :
Db 142 MVAGQVLDMGEQ 154

RESULT 10

US-09-489-039A-11571

; Sequence 11571, Application US/09489039A

```
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11571
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11571
```

```
Query Match          55.4%; Score 36; DB 4; Length 245;
Best Local Similarity 63.6%; Pred. No. 61;
Matches      7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 MYATEVLDLDG 11
        :: |||| |||
Db      72 VHGTEVLTLDG 82
```

RESULT 11

```
US-09-489-039A-8478
; Sequence 8478, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8478
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8478
```

```
Query Match          55.4%; Score 36; DB 4; Length 247;
Best Local Similarity 46.2%; Pred. No. 61;
Matches      6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 MYATEVLDLGSK 13
        :||| |::|:|
Db     112 VYATTVKEMEGNK 124
```

RESULT 12

US-09-512-251A-2

; Sequence 2, Application US/09512251A
 ; Patent No. 6555355
 ; GENERAL INFORMATION:
 ; APPLICANT: Hansen, Peter
 ; APPLICANT: Bauditz, Peter
 ; APPLICANT: Mikkelsen, Frank
 ; APPLICANT: Andersen, Kim
 ; TITLE OF INVENTION: Protease Variants and Compositions
 ; FILE REFERENCE: 5349.204-US
 ; CURRENT APPLICATION NUMBER: US/09/512,251A
 ; CURRENT FILING DATE: 2000-02-24
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 268
 ; TYPE: PRT
 ; ORGANISM: Bacillus

US-09-512-251A-2

Query Match 55.4%; Score 36; DB 4; Length 268;
 Best Local Similarity 58.3%; Pred. No. 67;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDGS 12
 :|| :||| :||
 Db 87 LYALKVLDNRGS 98

RESULT 13

US-09-515-150A-2

; Sequence 2, Application US/09515150A
 ; Patent No. 6558938
 ; GENERAL INFORMATION:
 ; APPLICANT: Hansen, Peter
 ; APPLICANT: Bauditz, Peter
 ; APPLICANT: Mikkelsen, Frank
 ; APPLICANT: Andersen, Kim
 ; TITLE OF INVENTION: Protease Variants and Compositions
 ; FILE REFERENCE: 5348.204-US
 ; CURRENT APPLICATION NUMBER: US/09/515,150A
 ; CURRENT FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 268
 ; TYPE: PRT
 ; ORGANISM: Bacillus

US-09-515-150A-2

Query Match 55.4%; Score 36; DB 4; Length 268;
 Best Local Similarity 58.3%; Pred. No. 67;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDGS 12
 :|| :||| :||

Db 87 LYALKVLDNRGS 98

RESULT 14

US-09-196-281-5
; Sequence 5, Application US/09196281A
; Patent No. 6605458
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants And Compositions
; FILE REFERENCE: 5435.200-US
; CURRENT APPLICATION NUMBER: US/09/196,281A
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: 1332/97
; EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Bacillus
US-09-196-281-5

Query Match 55.4%; Score 36; DB 4; Length 268;
Best Local Similarity 58.3%; Pred. No. 67;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDGS 12
: || : ||| : ||
Db 87 LYALKVLDNRGS 98

RESULT 15

US-10-007-389-2
; Sequence 2, Application US/10007389
; Patent No. 6727067
; GENERAL INFORMATION:
; APPLICANT: Russman, Eberhard
; APPLICANT: Meier, Thomas
; APPLICANT: Schmuck, Ranier
; APPLICANT: Staepels, Johnny
; APPLICANT: Wehnes, Uwe
; TITLE OF INVENTION: Methods for the analysis of non-proteinaceous
; TITLE OF INVENTION: components using a protease from a Bacillus strain
; FILE REFERENCE: Esperase
; CURRENT APPLICATION NUMBER: US/10/007,389
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Bacillus lentus
US-10-007-389-2

Query Match 55.4%; Score 36; DB 4; Length 268;
Best Local Similarity 58.3%; Pred. No. 67;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGS 12
: || : ||| : ||
Db 87 LYALKVLDNRGS 98

Search completed: February 10, 2005, 16:02:09
Job time : 30.1127 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:49:10 ; Search time 77.8169 Seconds
(without alignments)
54.586 Million cell updates/sec

Title: US-10-067-484-7
Perfect score: 65
Sequence: 1 MYATEVLDDLGSK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	65	100.0	13	14	US-10-067-484-7	Sequence 7, Appli
2	65	100.0	13	14	US-10-067-620-7	Sequence 7, Appli
3	39	60.0	287	16	US-10-437-963-146655	Sequence 146655,
4	39	60.0	299	16	US-10-767-701-36797	Sequence 36797, A
5	39	60.0	309	15	US-10-425-114-65406	Sequence 65406, A
6	39	60.0	721	15	US-10-094-749-3065	Sequence 3065, Ap
7	38	58.5	77	15	US-10-335-977-6260	Sequence 6260, Ap
8	38	58.5	170	16	US-10-767-701-32427	Sequence 32427, A
9	38	58.5	259	15	US-10-335-977-8681	Sequence 8681, Ap
10	38	58.5	286	15	US-10-276-774-1412	Sequence 1412, Ap
11	38	58.5	286	15	US-10-296-115-1340	Sequence 1340, Ap
12	38	58.5	286	15	US-10-296-115-1465	Sequence 1465, Ap
13	38	58.5	436	15	US-10-369-493-13460	Sequence 13460, A
14	38	58.5	436	15	US-10-369-493-18273	Sequence 18273, A
15	38	58.5	449	15	US-10-335-977-6263	Sequence 6263, Ap
16	38	58.5	457	14	US-10-156-761-11073	Sequence 11073, A
17	38	58.5	546	15	US-10-264-237-2780	Sequence 2780, Ap
18	38	58.5	600	15	US-10-015-115-73	Sequence 73, Appl
19	38	58.5	704	9	US-09-738-973-191	Sequence 191, App
20	38	58.5	704	9	US-09-854-133-191	Sequence 191, App
21	38	58.5	704	14	US-10-144-649A-191	Sequence 191, App
22	37	56.9	143	16	US-10-767-701-52827	Sequence 52827, A
23	37	56.9	162	10	US-09-765-815-14	Sequence 14, Appl
24	37	56.9	268	14	US-10-313-853-2	Sequence 2, Appli
25	37	56.9	272	16	US-10-872-166-9	Sequence 9, Appli
26	37	56.9	272	17	US-10-873-917-9	Sequence 9, Appli
27	37	56.9	280	15	US-10-264-213-206	Sequence 206, App
28	37	56.9	283	15	US-10-264-213-240	Sequence 240, App
29	37	56.9	375	10	US-09-813-408-8	Sequence 8, Appli
30	37	56.9	408	14	US-10-156-761-14852	Sequence 14852, A
31	37	56.9	456	15	US-10-424-599-178453	Sequence 178453,
32	37	56.9	667	16	US-10-437-963-202086	Sequence 202086,
33	37	56.9	948	15	US-10-282-122A-65876	Sequence 65876, A
34	37	56.9	982	15	US-10-282-122A-65306	Sequence 65306, A
35	36	55.4	13	16	US-10-468-496-913	Sequence 913, App
36	36	55.4	76	15	US-10-424-599-154529	Sequence 154529,
37	36	55.4	85	15	US-10-424-599-244166	Sequence 244166,
38	36	55.4	90	15	US-10-424-599-154528	Sequence 154528,
39	36	55.4	167	15	US-10-335-977-8928	Sequence 8928, Ap
40	36	55.4	181	15	US-10-335-977-8929	Sequence 8929, Ap
41	36	55.4	187	15	US-10-425-114-36598	Sequence 36598, A
42	36	55.4	266	9	US-09-837-235-18	Sequence 18, Appl
43	36	55.4	267	10	US-09-813-408-22	Sequence 22, Appl

44	36	55.4	268	14	US-10-007-389-2
45	36	55.4	268	14	US-10-336-324-2

Sequence 2, Appli
Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-067-484-7

```
; Sequence 7, Application US/10067484
; Publication No. US20030170763A1
; GENERAL INFORMATION:
; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; TITLE OF INVENTION: RAGWEED ALLERGENS
; FILE REFERENCE: 416272000200
; CURRENT APPLICATION NUMBER: US/10/067,484
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Ragweed
US-10-067-484-7
```

```
Query Match          100.0%; Score 65; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MYATEVLDLDGSK 13
          |||||||||||
Db      1 MYATEVLDLDGSK 13
```

RESULT 2

US-10-067-620-7

```
; Sequence 7, Application US/10067620
; Publication No. US20030180225A1
; GENERAL INFORMATION:
; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; APPLICANT: Teuber, Suzanne S.
; TITLE OF INVENTION: WALNUT AND RYEGRASS ALLERGENS
; FILE REFERENCE: 416272003400
; CURRENT APPLICATION NUMBER: US/10/067,620
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 13
```

; TYPE: PRT
; ORGANISM: Ragweed
US-10-067-620-7

Query Match 100.0%; Score 65; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYATEVLDLGSK 13
|||
Db 1 MYATEVLDLGSK 13

RESULT 3

US-10-437-963-146655
; Sequence 146655, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 146655
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(287)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_47259C.1.pep
US-10-437-963-146655

Query Match 60.0%; Score 39; DB 16; Length 287;
Best Local Similarity 58.3%; Pred. No. 91;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YATEVLDLGSK 13
: |||: ||| |:
Db 261 FGTEVVDLSSE 272

RESULT 4

US-10-767-701-36797
; Sequence 36797, Application US/10767701

; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 36797
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(299)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C11115_1.pep
US-10-767-701-36797

Query Match 60.0%; Score 39; DB 16; Length 299;
Best Local Similarity 53.8%; Pred. No. 95;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDLGSK 13
||| | : ||| :
Db 287 MYADEFITLDGNR 299

RESULT 5

US-10-425-114-65406
; Sequence 65406, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65406
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4763-011-D12_FLI.pep

US-10-425-114-65406

Query Match 60.0%; Score 39; DB 15; Length 309;
Best Local Similarity 53.8%; Pred. No. 98;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDGSK 13
||| | : |||::
Db 297 MYADEFMTLDGNR 309

RESULT 6

US-10-094-749-3065

; Sequence 3065, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3065
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-3065

Query Match 60.0%; Score 39; DB 15; Length 721;
Best Local Similarity 70.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YATEVLDLDG 11
| :||:||||
Db 154 YRSEVVDLDG 163

RESULT 7

US-10-335-977-6260

; Sequence 6260, Application US/10335977

; Publication No. US20040052799A1

; GENERAL INFORMATION:

; APPLICANT: DOUGLAS SMITH et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

; RELATING TO HELICOBACTER PYLORI FOR

; DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 10031

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: Windows NT 4.0

; SOFTWARE: UNIX

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/335,977

; FILING DATE: 30-Dec-2002

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/993,002

; FILING DATE: 17-DEC-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Mandragouras, Amy E.

; REGISTRATION NUMBER: 36,207

; REFERENCE/DOCKET NUMBER: GTN-018

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)227-7400

; TELEFAX: (617)742-4214

; INFORMATION FOR SEQ ID NO: 6260:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 77 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Helicobacter pylori

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (B) LOCATION 1...77

; SEQUENCE DESCRIPTION: SEQ ID NO: 6260:

US-10-335-977-6260

Query Match 58.5%; Score 38; DB 15; Length 77;

Best Local Similarity 58.3%; Pred. No. 33;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy

2 YATEVLDLGSK 13

Db 50 YAFEVLSVDGAR 61

RESULT 8

US-10-767-701-32427
; Sequence 32427, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 32427
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C12971_1.pep
US-10-767-701-32427

Query Match 58.5%; Score 38; DB 16; Length 170;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 YATEVLDLD 10
|||::|||
Db 110 YATQIVDLD 118

RESULT 9

US-10-335-977-8681
; Sequence 8681, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0

```

; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 8681:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 259 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...259
; SEQUENCE DESCRIPTION: SEQ ID NO: 8681:
US-10-335-977-8681

```

```

Query Match      58.5%; Score 38; DB 15; Length 259;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches      7; Conservative      3; Mismatches      2; Indels      0; Gaps      0;

```

```

Qy      2 YATEVLDLDGSK 13
      || ||| :||::
Db      232 YAFEVLSVDGAR 243

```

RESULT 10

```

US-10-276-774-1412
; Sequence 1412, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al.
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1412
; LENGTH: 286

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-1412

Query Match 58.5%; Score 38; DB 15; Length 286;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGGSK 13
:| :| |||||
Db 202 LYLKDVQDLGGK 214

RESULT 11

US-10-296-115-1340
; Sequence 1340, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1340
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1340

Query Match 58.5%; Score 38; DB 15; Length 286;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGGSK 13
:| :| |||||
Db 202 LYLKDVQDLGGK 214

RESULT 12

US-10-296-115-1465
; Sequence 1465, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1465
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1465

Query Match 58.5%; Score 38; DB 15; Length 286;
Best Local Similarity 53.8%; Pred.. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MYATEVLDLGSK 13
:| :| ||| |
Db 202 LYLKDVQDLGGK 214

RESULT 13

US-10-369-493-13460

; Sequence 13460, Application US/10369493
; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION
OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13460
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Thermoplasma volcanium
US-10-369-493-13460

Query Match 58.5%; Score 38; DB 15; Length 436;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDDL 10
|| |||||
Db 146 MYNREVLDDL 155

RESULT 14

US-10-369-493-18273

; Sequence 18273, Application US/10369493
; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION
OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18273
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Thermoplasma acidophilum
US-10-369-493-18273

Query Match 58.5%; Score 38; DB 15; Length 436;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDDL 10
|| |||||
Db 146 MYNREVLDDL 155

RESULT 15

US-10-335-977-6263

; Sequence 6263, Application US/10335977

; Publication No. US20040052799A1

; GENERAL INFORMATION:

; APPLICANT: DOUGLAS SMITH et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

; RELATING TO HELICOBACTER PYLORI FOR

; DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 10031

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD/ROM ISO9660

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: Windows NT 4.0

; SOFTWARE: UNIX

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/335,977

; FILING DATE: 30-Dec-2002

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/993,002

; FILING DATE: 17-DEC-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Mandragouras, Amy E.

; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6263:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...449
; SEQUENCE DESCRIPTION: SEQ ID NO: 6263:
US-10-335-977-6263

Query Match 58.5%; Score 38; DB 15; Length 449;
Best Local Similarity 58.3%; Pred. No. 2.2e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVLDLGSK 13
|| ||| :||::
Db 422 YAFEVLSVDGAR 433

Search completed: February 10, 2005, 16:41:33
Job time : 78.8169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 20.1408 Seconds
(without alignments)
62.104 Million cell updates/sec

Title: US-10-067-484-7
Perfect score: 65
Sequence: 1 MYATEVLDLGSK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	41	63.1	334	2	B84432	hypothetical prote
2	41	63.1	1376	2	S63986	collagen alpha 5 c
3	38	58.5	441	2	B71816	hypothetical prote
4	38	58.5	466	2	T44350	hypothetical prote
5	37	56.9	203	2	A45463	glutathione transf
6	37	56.9	319	2	S73159	hypothetical prote
7	37	56.9	361	2	A48373	high-alkaline seri
8	37	56.9	361	2	G83756	subtilisin-type al
9	37	56.9	370	2	AD2375	hypothetical prote
10	37	56.9	374	2	I39781	subtilisin (EC 3.4
11	37	56.9	433	2	T44587	cytochrome P450 ho
12	37	56.9	513	1	A35742	aqualysin (EC 3.4.
13	37	56.9	948	2	B81883	excinnuclease ABC c
14	37	56.9	949	2	A81138	excinnuclease ABC c
15	37	56.9	1478	2	S78131	DNA-directed RNA p
16	37	56.9	1777	2	T34369	hypothetical prote
17	36	55.4	165	2	D64648	hypothetical prote
18	36	55.4	167	2	B71939	hypothetical prote
19	36	55.4	198	2	S55131	hypothetical prote
20	36	55.4	215	2	B35534	hypothetical 23K p
21	36	55.4	275	2	AI1447	gp17 (Bacteriophag
22	36	55.4	281	2	C82102	conserved hypothet
23	36	55.4	305	2	F86744	tagatose-6-phospha
24	36	55.4	343	2	T36891	hypothetical prote
25	36	55.4	382	1	SUBSN	subtilisin (EC 3.4
26	36	55.4	388	2	D84992	hypothetical prote
27	36	55.4	398	2	AE3577	sugar-binding prot
28	36	55.4	408	2	S76678	hypothetical prote
29	36	55.4	416	1	H64328	hypothetical prote
30	36	55.4	417	2	T47767	hypothetical prote
31	36	55.4	433	2	AI2954	hypothetical prote
32	36	55.4	433	2	D98328	hypothetical ABC t
33	36	55.4	436	2	T24953	hypothetical prote
34	36	55.4	453	2	S53995	TodX protein - Pse
35	36	55.4	647	1	ESECPC	2',3'-cyclic-nucle
36	36	55.4	647	1	G86118	2',3'-cyclic-nucle
37	36	55.4	647	1	G91277	2',3'-cyclic-nucle
38	36	55.4	733	2	AD2444	hypothetical prote
39	36	55.4	736	2	S45859	hypothetical prote

40	36	55.4	755	2	T46411	hypothetical prote
41	36	55.4	829	2	E87305	TonB-dependent rec
42	36	55.4	906	2	G96621	probable disease r
43	36	55.4	1048	2	T23764	hypothetical prote
44	36	55.4	1133	1	EGRT	epidermal growth f
45	36	55.4	1207	2	B88789	protein ZK1251.9 [

ALIGNMENTS

RESULT 1

B84432

hypothetical protein At2g02030 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: B84432

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhagen, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: B84432

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-334 <STO>

A;Cross-references: UNIPROT:Q9ZPS1; GB:AE002093; NID:g4406785; PIDN:AAD20095.1; GSPDB:GN00139.

C;Genetics:

A;Gene: At2g02030

A;Map position: 2

Query Match 63.1%; Score 41; DB 2; Length 334;
 Best Local Similarity 52.4%; Pred. No. 9.6;
 Matches 11; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

Qy	1	MYAT-----EVL	DL	DGSK	13
Db	220	MYNTSPATPPTCEV	LD	LDGKK	240

RESULT 2

S63986

collagen alpha 5 chain - sea urchin (Strongylocentrotus purpuratus) (fragment)

C;Species: Strongylocentrotus purpuratus (purple urchin)

C;Date: 20-Jul-1996 #sequence_revision 08-Nov-1996 #text_change 09-Jul-2004

C;Accession: S63986; S64638

R;Exposito, J.Y.; Boute, N.; Deleage, G.; Garrone, R.

Eur. J. Biochem. 234, 59-65, 1995

A;Title: Characterization of two genes coding for a similar four-cysteine motif of the amino-terminal propeptide of a sea urchin fibrillar collagen.

A;Reference number: S63985; MUID:96096722; PMID:8529669
 A;Accession: S63986
 A;Status: nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 1-1376 <EXP>
 A;Cross-references: UNIPROT:Q26637; EMBL:X89804
 R;Exposito, J.Y.
 submitted to the EMBL Data Library, July 1995
 A;Reference number: S64637
 A;Accession: S64638
 A;Molecule type: DNA
 A;Residues: 1-658, 'G', 660-870, 'G', 872-901, 'H', 903-1185, 'T', 1187-1214, 'Y', 1216-1376 <EXW>
 A;Cross-references: EMBL:X89804
 C;Genetics:
 A;Gene: COLP5alpha
 A;Introns: 73/1; 136/2; 221/1; 369/1; 517/1; 659/1; 799/1; 948/1; 1093/1; 1236/1
 C;Keywords: extracellular matrix
 F;15-73/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 63.1%; Score 41; DB 2; Length 1376;
 Best Local Similarity 63.6%; Pred. No. 44;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVLDDLGS 12
 | ::|||||
 Db 580 YVEQILDDLGS 590

RESULT 3

B71816

hypothetical protein jhp1383 - *Helicobacter pylori* (strain J99)

C;Species: *Helicobacter pylori*

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C;Accession: B71816

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.;
 Smith, D.R.; Noonan, B.; Guild, B.C.; deJonge, B.L.; Carmel, G.; Tummino, P.J.;
 Caruso, A.; Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.; Merberg, D.;
 Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Trust, T.J.
 Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*.

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: B71816

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-441 <ARN>

A;Cross-references: UNIPROT:Q9ZJC9; GB:AE001560; GB:AE001439; NID:g4155981;
 PIDN:AAD06954.1; PID:g4155992

A;Experimental source: strain J99

C;Genetics:

A;Gene: jhp1383

C;Superfamily: hypothetical protein HI0107

Query Match 58.5%; Score 38; DB 2; Length 441;
 Best Local Similarity 58.3%; Pred. No. 46;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy      2 YATEVLDLDGSK 13
        || ||| :||:
Db      414 YAFEVLSVDGAR 425
```

RESULT 4

T44350

hypothetical protein [imported] - Clostridium histolyticum

C;Species: Clostridium histolyticum

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C;Accession: T44350

R;Matsushita, O.; Jung, C.M.; Katayama, S.; Minami, J.; Takahashi, Y.; Okabe, A.
J. Bacteriol. 181, 923-933, 1999

A;Title: Gene duplication and multiplicity of collagenases in Clostridium
histolyticum.

A;Reference number: Z22752; MUID:99121032; PMID:9922257

A;Accession: T44350

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-466 <MAT>

A;Cross-references: UNIPROT:Q9ZNK3; EMBL:AB014075; NID:g3868863;

PIDN:BAA34256.1; PID:g3868868

A;Experimental source: strain JCM 1403

C;Superfamily: hypothetical protein b1439

Query Match 58.5%; Score 38; DB 2; Length 466;

Best Local Similarity 88.9%; Pred. No. 49;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      5 EVLDLDGSK 13
        ||| |||
Db      129 EVLDRDGSK 137
```

RESULT 5

A45463

glutathione transferase (EC 2.5.1.18) - Sloane's squid

C;Species: Ommastrephes sloanei pacificus (Sloane's squid)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: A45463

R;Tomarev, S.I.; Zinovieva, R.D.; Guo, K.; Piatigorsky, J.
J. Biol. Chem. 268, 4534-4542, 1993

A;Title: Squid glutathione S-transferase. Relationships with other glutathione
S-transferases and S-crystallins of cephalopods.

A;Reference number: A45463; MUID:93179471; PMID:8440736

A;Accession: A45463

A;Molecule type: DNA; mRNA; protein

A;Residues: 1-203 <TOM>

A;Cross-references: UNIPROT:P46088; GB:L02054; NID:g159847; PIDN:AAA92066.1;

PID:g1223936

A;Experimental source: digestive gland

A;Note: sequence extracted from NCBI backbone (NCBIN:125992, NCBIP:125993)

A;Note: 124-Tyr was also found; enzyme activity was demonstrated

C;Superfamily: glutathione transferase

C;Keywords: glutathione; transferase

F;8/Active site: Tyr #status predicted
F;14,43/Binding site: substrate (Arg, Lys) #status predicted

Query Match 56.9%; Score 37; DB 2; Length 203;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

Qy 1 MY--ATEVLDLDGSK 13
|| | |||:|:|
Db 46 MYSNAMPVLDIDGTK 60

RESULT 6

S73159

hypothetical protein 39 - red alga (*Porphyra purpurea*) chloroplast

C;Species: chloroplast *Porphyra purpurea*

C;Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004

C;Accession: S73159

R;Reith, M.; Munholland, J.

Plant Mol. Biol. Rep. 13, 333-335, 1995

A;Title: Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome.

A;Reference number: S73108

A;Accession: S73159

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-319 <REI>

A;Cross-references: UNIPROT:P51238; EMBL:U38804; NID:g1276652; PID:g1276704

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C;Genetics:

A;Gene: ycf39

A;Genome: chloroplast

C;Keywords: chloroplast

Query Match 56.9%; Score 37; DB 2; Length 319;
Best Local Similarity 70.0%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVLDLDG 11
| || :|||
Db 80 YNTEQIDLDG 89

RESULT 7

A48373

high-alkaline serine proteinase (EC 3.4.21.-) precursor - *Bacillus* sp. (strain AH-101)

N;Alternate names: subtilisin-like thermostable alkaline serine proteinase

C;Species: *Bacillus* sp.

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 28-May-1999

C;Accession: A48373; JS0714

R;Takami, H.; Kobayashi, T.; Aono, R.; Horikoshi, K.

Appl. Microbiol. Biotechnol. 38, 101-108, 1992

A;Title: Molecular cloning, nucleotide sequence and expression of the structural gene for a thermostable alkaline protease from *Bacillus* sp. no. AH-101.

A;Reference number: A48373; MUID:93098926; PMID:1369007

A;Accession: A48373
 A;Molecule type: DNA
 A;Residues: 1-361 <TAK>
 A;Cross-references: GB:S50880; NID:g261737; PIDN:AAC60421.1; PID:g261738
 A;Experimental source: AH-101
 A;Note: this sequence is inconsistent with the nucleotide translation
 A;Note: sequence extracted from NCBI backbone (NCBIN:121090, NCBIP:121091)
 R;Takami, H.; Kobayashi, T.; Aono, R.; Horikoshi, K.
 submitted to JIPID, July 1992
 A;Description: Molecular cloning, nucleotide sequence and expression of the
 structural gene for a thermostable alkaline protease from *Bacillus* sp. no. AH-
 101.
 A;Reference number: JS0714
 A;Accession: JS0714
 A;Molecule type: DNA
 A;Residues: 94-334, 'L', 336-361 <TA2>
 C;Comment: This alkaliphilic *Bacillus* homolog to the subtilisins of neutrophilic
Bacilli has a pH optimum of 12-13.
 C;Superfamily: subtilisin; subtilisin homology
 C;Keywords: extracellular protein; hydrolase; serine proteinase
 F;115-321/Domain: subtilisin homology <SBT>
 F;124,154,307/Active site: Asp, His, Ser #status predicted

Query Match 56.9%; Score 37; DB 2; Length 361;
 Best Local Similarity 58.3%; Pred. No. 57;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MYATEVLDLDGS 12
 :|| :||| :||
 Db 180 LYAVKVLDRNGS 191

RESULT 6

G83756

subtilisin-type alkaline proteinase (EC 3.4.21.-) BH0855 precursor [similarity]
 - *Bacillus halodurans* (strain C-125)

C;Species: *Bacillus halodurans*

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C;Accession: G83756

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji,
 F.; Hirama, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S.; Horikoshi, K.

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus*
halodurans and genomic sequence comparison with *Bacillus subtilis*.

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: G83756

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-361 <STO>

A;Cross-references: UNIPROT:P41363; GB:AP001510; GB:BA000004; NID:g10173440;
 PIDN:BAB04574.1; GSPDB:GN00137

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH0855

C;Superfamily: subtilisin; subtilisin homology

C;Keywords: hydrolase; serine proteinase

F;1-25/Domain: signal sequence #status predicted <SIG>

Query Match 56.9%; Score 37; DB 2; Length 361;
Best Local Similarity 58.3%; Pred. No. 57;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGS 12
:| | :| | | :| |
Db 180 LYAVKVLDRNGS 191

RESULT 9

AD2375

hypothetical protein all4556 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AD2375

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.;
Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.;
Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.;
Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing
Cyanobacterium Anabaena sp. strain PCC 7120.

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AD2375

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-370 <KUR>

A;Cross-references: UNIPROT:Q8YNK9; GB:BA000019; PIDN:BAB76255.1; PID:g17133692;
GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all4556

Query Match 56.9%; Score 37; DB 2; Length 370;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YATEVLDDLGD 11
:| | :| | :| |
Db 334 FAGEILDLIDG 343

RESULT 10

I39781

subtilisin (EC 3.4.21.62) ALP I precursor - Bacillus sp.

C;Species: Bacillus sp.

C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004

C;Accession: I39781

R;Yamagata, Y.; Sato, T.; Hanzawa, S.; Ichishima, E.

Curr. Microbiol. 30, 201-209, 1995

A;Title: The structure of subtilisin ALP I from alkalophilic Bacillus sp. NKS-
21.

A;Reference number: I39781; MUID:95195580; PMID:7765893

A;Accession: I39781

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA
A;Residues: 1-374 <RES>
A;Cross-references: UNIPROT:Q45523; GB:D29736; NID:g975628; PIDN:BAA06158.1;
PID:g975629
C;Genetics:
A;Gene: aprQ
C;Superfamily: subtilisin; subtilisin homology
C;Keywords: hydrolase; serine proteinase
F;124-334/Domain: subtilisin homology <SBT>

Query Match 56.9%; Score 37; DB 2; Length 374;
Best Local Similarity 58.3%; Pred. No. 59;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDGS 12
: || : ||| : ||
Db 189 LYAVKVLDRNGS 200

RESULT 11

T44587

cytochrome P450 homolog [imported] - *Streptomyces fradiae*

C;Species: *Streptomyces fradiae*

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C;Accession: T44587

R;Bate, N.; Butler, A.R.; Gandeck, A.R.; Cundliffe, E.

Chem. Biol. 6, 617-624, 1999

A;Title: Multiple regulatory genes in the tylosin-biosynthetic cluster of *Streptomyces fradiae*.

A;Reference number: Z22801; MUID:99398833; PMID:10467127

A;Accession: T44587

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-433 <BAT>

A;Cross-references: UNIPROT:Q9XCC6; EMBL:AF145049; PIDN:AAD40802.1

A;Experimental source: strain T59235

C;Superfamily: Bacillus cytochrome P450 CYP106; cytochrome P450 homology

F;262-399/Domain: cytochrome P450 homology <P45>

Query Match 56.9%; Score 37; DB 2; Length 433;
Best Local Similarity 70.0%; Pred. No. 69;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVLDLDG 11
|| | : |||
Db 319 YAVEDIDLGD 328

RESULT 12

A35742

aqualysin (EC 3.4.21.-) I precursor - *Thermus aquaticus*

C;Species: *Thermus aquaticus*

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: A35742; S00620; S00324

R;Terada, I.; Kwon, S.T.; Miyata, Y.; Matsuzawa, H.; Ohta, T.

J. Biol. Chem. 265, 6576-6581, 1990

A;Title: Unique precursor structure of an extracellular protease, aqualysin I, with NH-2- and COOH-terminal pro-sequences and its processing in *Escherichia coli*.
A;Reference number: A35742; MUID:90216674; PMID:2182621
A;Accession: A35742
A;Molecule type: DNA
A;Residues: 1-513 <TER>
A;Cross-references: UNIPROT:P08594; GB:J90108; GB:D90108; GB:J05414; NID:g217171; PIDN:BAA14135.1; PID:g217172
A;Note: the authors translated the codon CTG for residue 470 as Val, and GGT for residue 473 as Ala
R;Kwon, S.T.; Terada, I.; Matsuzawa, H.; Ohta, T.
Eur. J. Biochem. 173, 491-497, 1988
A;Title: Nucleotide sequence of the gene for aqualysin I (a thermophilic alkaline serine protease) of *Thermus aquaticus* YT-1 and characteristics of the deduced primary structure of the enzyme.
A;Reference number: S00620; MUID:88225062; PMID:3286255
A;Accession: S00620
A;Molecule type: DNA
A;Residues: 75-442 <KWO>
A;Cross-references: EMBL:X07734; NID:g48069; PIDN:CAA30559.1; PID:g602091
A;Note: part of this sequence, including the amino and carboxyl ends of the mature protein, was confirmed by protein sequencing
R;Matsuzawa, H.; Tokugawa, K.; Hamaoki, M.; Mizoguchi, M.; Taguchi, H.; Terada, I.; Kwon, S.T.; Ohta, T.
Eur. J. Biochem. 171, 441-447, 1988
A;Title: Purification and characterization of aqualysin I (a thermophilic alkaline serine protease) produced by *Thermus aquaticus* YT-1.
A;Reference number: S00324; MUID:88151937; PMID:3162211
A;Accession: S00324
A;Molecule type: protein
A;Residues: 128-170 <MATS>
C;Superfamily: subtilisin; subtilisin homology
C;Keywords: extracellular protein; hydrolase; serine proteinase
F;1-14/Domain: signal sequence #status predicted <SIG>
F;15-127/Domain: propeptide #status predicted <PRO>
F;128-408/Product: aqualysin I #status experimental <MAT>
F;157-364/Domain: subtilisin homology <SBT>
F;255-257,281-283/Region: S1 specificity crevice #status predicted
F;409-513/Domain: carboxyl-terminal propeptide #status predicted <CPR>
F;166,197,349/Active site: Asp, His, Ser #status predicted

Query Match 56.9%; Score 37; DB 1; Length 513;
Best Local Similarity 58.3%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGS 12
: || ||| : ||
Db 218 LYAVRVLDCNGS 229

RESULT 13

B81883

excinuclease ABC chain A NMA1159 [similarity] - *Neisseria meningitidis* (strain Z2491 serogroup A)

N;Contains: excision endonuclease ABC (EC 3.1.--.) chain A

C;Species: *Neisseria meningitidis*

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
 C;Accession: B81883
 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morelli, G.; Basham, D.; Brown, D.; Chillingworth, T.; Davies, R.M.; Davis, P.; Devlin, K.; Feltwell, T.; Hamlin, N.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, M.A.; Rutherford, K.M.; Simmonds, M.; Skelton, J.; Whitehead, S.; Spratt, B.G.; Barrell, B.G.
 Nature 404, 502-506, 2000
 A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A;Reference number: A81775; MUID:20222556; PMID:10761919
 A;Accession: B81883
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-948 <PAR>
 A;Cross-references: UNIPROT:Q9JUS4; GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84421.1; PID:g7379852; GSPDB:GN00124; NMAASP:NMA1159
 A;Experimental source: serogroup A, strain Z2491
 C;Genetics:
 A;Gene: uvrA; NMA1159
 C;Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
 C;Keywords: ATP; DNA binding; DNA repair; hydrolase; nucleotide binding; P-loop
 F;42-49/Region: nucleotide-binding motif A (P-loop)
 F;649-656/Region: nucleotide-binding motif A (P-loop)

Query Match 56.9%; Score 37; DB 2; Length 948;
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 EVLDLDGSK 13
 | | | | | | |
 Db 453 ETLDLDGNK 461

RESULT 14

A81138
 excinuclease ABC chain A NMB0962 [similarity] - *Neisseria meningitidis* (strain MC58 serogroup B)
 N;Contains: excision endonuclease ABC (EC 3.1.-.-) chain A
 C;Species: *Neisseria meningitidis*
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C;Accession: A81138
 R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Ketchum, K.A.; Hood, D.W.; Peden, J.F.; Dodson, R.J.; Nelson, W.C.; Gwinn, M.L.; DeBoy, R.; Peterson, J.D.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Mason, T.; Ciecko, A.; Parksey, D.S.; Blair, E.; Cittone, H.; Clark, E.B.; Cotton, M.D.; Utterback, T.R.; Khouri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, J.C.
 A;Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
 A;Reference number: A81000; MUID:20175755; PMID:10710307
 A;Accession: A81138
 A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-949 <TET>
A;Cross-references: UNIPROT:Q9JZP1; GB:AE002447; GB:AE002098; NID:g7226196;
PIDN:AAF41368.1; PID:g7226202; GSPDB:GN00119; TIGR:NMB0962
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0962
C;Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
C;Keywords: ATP; DNA binding; DNA repair; hydrolase; nucleotide binding; P-loop
F;42-49/Region: nucleotide-binding motif A (P-loop)
F;649-656/Region: nucleotide-binding motif A (P-loop)

Query Match 56.9%; Score 37; DB 2; Length 949;
Best Local Similarity 77.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 EVLDLDGSK 13
| | | | | : |
Db 453 ETLDLGDK 461

RESULT 15

S78131

DNA-directed RNA polymerase (EC 2.7.7.6) chain beta - *Reclinomonas americana*
(ATCC 50394) mitochondrion

C;Species: mitochondrion *Reclinomonas americana*

A;Variety: ATCC 50394

C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004

C;Accession: S78131

R;Lang, B.F.; Burger, G.; O'Kelly, C.J.; Cedergren, R.; Golding, G.B.; Lemieux, C.; Sankoff, D.; Turmel, M.; Gray, M.W.

Nature 387, 493-497, 1997

A;Title: An ancestral mitochondrial DNA resembling a eubacterial genome in miniature.

A;Reference number: S78127; MUID:97311393; PMID:9168110

A;Accession: S78131

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1478 <LAN>

A;Cross-references: UNIPROT:O21237; EMBL:AF007261; NID:g2258325;

PIDN:AAD11864.1; PID:g2258330

A;Experimental source: ATCC 50394

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1997

C;Genetics:

A;Gene: rpoB

A;Genome: mitochondrion

C;Superfamily: DNA-directed RNA polymerase beta chain

C;Keywords: mitochondrion; nucleotidyltransferase; transcription

Query Match 56.9%; Score 37; DB 2; Length 1478;
Best Local Similarity 72.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVLDLDGS 12
: | | | | | | |
Db 631 HATEVLKKDGS 641

Search completed: February 10, 2005, 15:59:35
Job time : 21.1408 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 94.8451 Seconds
(without alignments)
70.188 Million cell updates/sec

Title: US-10-067-484-7
Perfect score: 65
Sequence: 1 MYATEVLDDL D GSK.13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	42	64.6	227	2	Q6G1R0	Q6g1r0 bartonella
2	41	63.1	132	2	Q82ZG5	Q82zg5 enterococcu
3	41	63.1	334	2	Q9ZPS1	Q9zps1 arabidopsis
4	41	63.1	1376	2	Q26637	Q26637 strongyloce
5	40	61.5	34	2	Q88FC8	Q88fc8 pseudomonas
6	40	61.5	350	2	Q92ST7	Q92st7 rhizobium m
7	40	61.5	567	2	Q7S1S4	Q7s1s4 neurospora
8	40	61.5	625	1	ACD9_MOUSE	Q8jzn5 mus musculu
9	39	60.0	221	2	Q63BX2	Q63bx2 bacillus ce
10	39	60.0	221	2	Q738Z3	Q738z3 bacillus ce
11	39	60.0	221	2	Q81R38	Q81r38 bacillus an

12	39	60.0	221	2	Q6HJE3	Q6hje3	bacillus th
13	39	60.0	315	2	Q8H9W4	Q8h9w4	pseudomonas
14	39	60.0	362	2	Q88CQ1	Q88cq1	pseudomonas
15	39	60.0	391	2	Q6CSK6	Q6csk6	kluyveromyc
16	39	60.0	395	2	Q971Y8	Q971y8	sulfolobus
17	39	60.0	433	2	Q747F9	Q747f9	geobacter s
18	39	60.0	488	2	Q6C2H5	Q6c2h5	yarrowia li
19	39	60.0	721	2	Q96M80	Q96m80	homo sapien
20	39	60.0	1235	2	Q86YZ7	Q86yz7	homo sapien
21	39	60.0	1260	2	Q86YZ8	Q86yz8	homo sapien
22	39	60.0	1308	1	CTA4_HUMAN	Q9c0a0	homo sapien
23	39	60.0	1311	2	Q8WX98	Q8wx98	homo sapien
24	39	60.0	1401	2	Q7QXR1	Q7qxr1	giardia lam
25	38	58.5	155	2	Q6A2Q8	Q6a2q8	populus tre
26	38	58.5	155	2	Q6A2R3	Q6a2r3	populus tre
27	38	58.5	172	2	Q6A2M6	Q6a2m6	populus tre
28	38	58.5	172	2	Q6A2N8	Q6a2n8	populus tre
29	38	58.5	172	2	Q6A2P4	Q6a2p4	populus tre
30	38	58.5	172	2	Q6A2Q4	Q6a2q4	populus tre
31	38	58.5	172	2	Q8GTC3	Q8gtc3	populus tre
32	38	58.5	197	2	Q6N1D5	Q6n1d5	rhodopseudo
33	38	58.5	218	2	Q6FWW1	Q6fww1	candida gla
34	38	58.5	222	2	Q7QGS6	Q7qgs6	anopheles g
35	38	58.5	287	1	TRUB_FUSNN	Q8r5x8	fusobacteri
36	38	58.5	353	2	Q8ET00	Q8et00	oceanobacil
37	38	58.5	356	2	Q7NC80	Q7nc80	gloeobacter
38	38	58.5	389	2	Q8PTR4	Q8ptr4	methanosarc
39	38	58.5	389	2	Q8TQ40	Q8tq40	methanosarc
40	38	58.5	436	2	Q97CK3	Q97ck3	thermoplasm
41	38	58.5	436	2	Q9HIA6	Q9hia6	thermoplasm
42	38	58.5	441	2	Q9ZJC9	Q9zjc9	helicobacte
43	38	58.5	457	2	Q82HH2	Q82hh2	streptomyce
44	38	58.5	466	2	Q9ZNK3	Q9znk3	clostridium
45	38	58.5	500	2	Q7PWN3	Q7pwn3	anopheles g

ALIGNMENTS

RESULT 1

Q6G1R0

ID Q6G1R0 PRELIMINARY; PRT; 227 AA.

AC Q6G1R0;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Two-component response regulator.

GN OrderedLocusNames=BH16140;

OS Bartonella henselae (Rochalimaea henselae).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bartonellaceae; Bartonella.

OX NCBI_TaxID=38323;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 49882 / Houston 1;

RX PubMed=15210978; DOI=10.1073/pnas.0305659101;

RA Alsmark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,

RA Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M.,
 RA La Scola B., Holmberg M., Andersson S.G.E.;
 RT "The louse-borne human pathogen Bartonella quintana is a genomic
 RT derivative of the zoonotic agent Bartonella henselae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721(2004).
 CC -!- SIMILARITY: Contains 1 response regulatory domain.
 DR EMBL; BX897699; CAF28377.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0000156; F:two-component response regulator activity; IEA.
 DR GO; GO:0007600; P:sensory perception; IEA.
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
 DR InterPro; IPR009059; bi_resp_regltr_C.
 DR InterPro; IPR011006; CheY_like.
 DR InterPro; IPR001789; Response_reg.
 DR InterPro; IPR001867; Trans_reg_C.
 DR Pfam; PF00072; Response_reg; 1.
 DR Pfam; PF00486; Trans_reg_C; 1.
 DR ProDom; PD000039; Response_reg; 1.
 DR ProDom; PD000329; Trans_reg_C; 1.
 DR SMART; SM00448; REC; 1.
 DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
 KW Complete proteome; DNA-binding; Phosphorylation; Sensory transduction;
 KW Transcription; Transcription regulation.
 SQ SEQUENCE 227 AA; 26329 MW; F1539D3C5A09BF35 CRC64;

Query Match 64.6%; Score 42; DB 2; Length 227;
 Best Local Similarity 61.5%; Pred. No. 20;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MYATEVLDLGSK 13
 ::|||: ||| |
 Db 53 IFATELPDLGDK 65

RESULT 2

Q82ZG5

ID Q82ZG5 PRELIMINARY; PRT; 132 AA.
 AC Q82ZG5;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glyoxalase family protein.
 GN OrderedLocusNames=EF3092;
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
 RA Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
 RA Fraser C.M.;

RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis.";
RL Science 299:2071-2074(2003).
DR EMBL; AE016956; AAO82773.1; -.
DR TIGR; EF3092; -.
DR InterPro; IPR004360; Gly_bleo_diox.
DR Pfam; PF00903; Glyoxalase; 1.
KW Complete proteome.
SQ SEQUENCE 132 AA; 15038 MW; 58E4BE3412BB384B CRC64;

Query Match 63.1%; Score 41; DB 2; Length 132;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGS 12
|| || ||||:
Db 112 MYGLEVQDL DGN 123

RESULT 3

Q9ZPS1

ID Q9ZPS1 PRELIMINARY; PRT; 334 AA.
AC Q9ZPS1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein At2g02030.
GN Name=At2g02030;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC006532; AAD20095.1; -.
DR PIR; B84432; B84432.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR011043; Gal_oxid_central.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PS50181; FBOX; 1.
KW Hypothetical protein.
SQ SEQUENCE 334 AA; 37978 MW; 0D827279DD2EEF3F CRC64;

Query Match 63.1%; Score 41; DB 2; Length 334;
Best Local Similarity 52.4%; Pred. No. 46;
Matches 11; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

Qy 1 MYAT-----EVLDDLGSK 13
 ||| |||||
 Db 220 MYNTSPATPPTCEVLDDLGGK 240

RESULT 4

Q26637

ID Q26637 PRELIMINARY; PRT; 1376 AA.
 AC Q26637;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE 5 alpha fibrillar collagen (Fragment).
 GN Name=COL5alpha;
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96096722; PubMed=8529669;
 RA Exposito J.Y., Boute N., Deleage G., Garrone R.;
 RT "Characterization of two genes coding for a similar four-cysteine
 RT motif of the amino-terminal propeptide of a sea urchin fibrillar
 RT collagen.";
 RL Eur. J. Biochem. 234:59-65(1995).
 DR EMBL; X89800; CAA61928.1; -.
 DR EMBL; X89801; CAA61928.1; JOINED.
 DR EMBL; X89802; CAA61928.1; JOINED.
 DR EMBL; X89803; CAA61928.1; JOINED.
 DR EMBL; X89804; CAA61928.1; JOINED.
 DR EMBL; X89805; CAA61928.1; JOINED.
 DR PIR; S63986; S63986.
 DR InterPro; IPR009041; PMP_SGCI.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00093; VWC; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
 DR PROSITE; PS50184; VWFC_2; 1.
 KW Collagen.
 FT NON_TER 1376 1376
 SQ SEQUENCE 1376 AA; 151182 MW; AF134036781FAAC6 CRC64;

Query Match 63.1%; Score 41; DB 2; Length 1376;
 Best Local Similarity 63.6%; Pred. No. 2e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVLDDLGS 12
 | ::|||
 Db 580 YVEQILDDLGS 590

RESULT 5

Q88FC8

ID Q88FC8 PRELIMINARY; PRT; 34 AA.

AC Q88FC8;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=PP4170;
 OS *Pseudomonas putida* (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; *Pseudomonas*.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
 RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
 RA Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
 RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile *Pseudomonas putida* KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL; AE016789; AAN69751.1; -.
 DR TIGR; PP4170; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 34 AA; 3652 MW; F9A1DF5546D81A06 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 34;
 Best Local Similarity 66.7%; Pred. No. 6.8;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 YATEVLDLDGSK 13
 | |||||:
 Db 9 YRVAVLDLDGSE 20

RESULT 6

Q92ST7

ID Q92ST7 PRELIMINARY; PRT; 350 AA.
 AC Q92ST7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE PUTATIVE LYSYL-TRNA SYNTHETASE PROTEIN (EC 6.1.1.6).
 GN ORFNames=SMc00356;
 OS *Rhizobium meliloti* (*Sinorhizobium meliloti*).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; *Sinorhizobium*/Ensifer group; *Sinorhizobium*.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT *Sinorhizobium meliloti* strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 DR EMBL; AL591783; CAC41713.1; -.
 DR HSSP; P13030; 1BBW.
 DR GO; GO:0005737; C:cytoplasm; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0004824; F:lysine-tRNA ligase activity; IEA.
 DR GO; GO:0006430; P:lysyl-tRNA aminoacylation; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR004364; tRNA-synt_2.
 DR InterPro; IPR002313; tRNA-synt_lys_2.
 DR InterPro; IPR006195; tRNA_ligase_II.
 DR Pfam; PF00152; tRNA-synt_2; 1.
 DR PRINTS; PR00982; TRNASYNTHLYS.
 DR PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
 KW ATP-binding; Aminoacyl-tRNA synthetase; Complete proteome; Ligase;
 KW Protein biosynthesis.
 SQ SEQUENCE 350 AA; 38718 MW; 265C42CDFB716B87 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 350;
 Best Local Similarity 72.7%; Pred. No. 75;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YATEVLDLDGS 12
 :||| | |||
 Db 60 FATEALGLDGS 70

RESULT 7

Q7S1S4

ID Q7S1S4 PRELIMINARY; PRT; 567 AA.
 AC Q7S1S4;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Name=NCU07754.1;
 OS *Neurospora crassa*.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; *Neurospora*.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OR74A;
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,

RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
 RA Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
 RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
 RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
 RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
 RT "The Genome Sequence of the Filamentous Fungus *Neurospora crassa*.";
 RL Nature 0:0-0(2003).
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AABX01000441; EAA29311.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR002293; AA/rel_permease1.
 DR InterPro; IPR004841; Permease_region.
 DR Pfam; PF00324; AA_permease; 1.
 KW Hypothetical protein; Transmembrane; Transport.
 SQ SEQUENCE 567 AA; 61017 MW; EBAE19DF138F19DE CRC64;

Query Match 61.5%; Score 40; DB 2; Length 567;
 Best Local Similarity 70.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ATEVLDLGDS 12
 |:|:|:|:
 Db 20 ASEILDVDGS 29

RESULT 8

ACD9_MOUSE

ID ACD9_MOUSE STANDARD; PRT; 625 AA.
 AC Q8JZN5; Q8BK76; Q8C0B5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Acyl-CoA dehydrogenase family member 9, mitochondrial precursor
 DE (EC 1.3.99.-) (ACAD-9).
 GN Name=Acad9;
 OS *Mus musculus* (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perteau G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Has a dehydrogenase activity on palmitoyl-CoA (C16:0)
 CC and stearoyl-CoA (C18:0). It is three times more active on
 CC palmitoyl-CoA than on stearoyl-CoA. Has little activity on
 CC octanoyl-CoA (C8:0), butyryl-CoA (C4:0) or isovaleryl-CoA (5:0)
 CC (By similarity).
 CC -!- COFACTOR: FAD (By similarity).
 CC -!- SUBCELLULAR LOCATION: Mitochondrial (Potential).
 CC -!- SIMILARITY: Belongs to the acyl-CoA dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AK031820; BAC27565.1; -.
 DR EMBL; AK075984; BAC36096.1; -.
 DR EMBL; BC031137; AAH31137.1; -.
 DR EMBL; BC032213; AAH32213.1; -.
 DR EMBL; BC033277; AAH33277.1; -.
 DR PIR; PT0697; PT0697.
 DR PIR; PT0721; PT0721.
 DR HSSP; P15651; 1JQI.
 DR MGD; MGI:1914272; Acad9.
 DR InterPro; IPR006089; Acyl-CoA_dh.
 DR InterPro; IPR006090; Acyl-CoA_dh_C.
 DR InterPro; IPR006091; Acyl-CoA_dh_M.
 DR InterPro; IPR006092; Acyl-CoA_dh_N.
 DR InterPro; IPR009100; AcylCoA_dehyd_NM.
 DR InterPro; IPR009075; AcylCoADH_C_like.
 DR Pfam; PF00441; Acyl-CoA_dh; 1.
 DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
 DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
 DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
 DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
 KW FAD; Flavoprotein; Mitochondrion; Oxidoreductase; Transit peptide.
 FT TRANSIT 1 ? Mitochondrion (Potential).
 FT CHAIN ? 625 Acyl-CoA dehydrogenase family member 9.
 FT ACT_SITE 430 430 Proton acceptor (By similarity).
 FT CONFLICT 15 15 A -> G (in Ref. 1; BAC27565).
 FT CONFLICT 53 53 K -> E (in Ref. 1).
 FT CONFLICT 163 163 D -> E (in Ref. 1).
 FT CONFLICT 540 540 I -> V (in Ref. 1; BAC27565).
 SQ SEQUENCE 625 AA; 68707 MW; 4F06FFFBFD82F022 CRC64;

Query Match 61.5%; Score 40; DB 1; Length 625;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 TEVLDLDGSK 13
 |||:| |||
 Db 233 TEVVDSGSK 242

RESULT 9

Q63BX2

ID Q63BX2 PRELIMINARY; PRT; 221 AA.
 AC Q63BX2;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Phosphoglycolate phosphatase (EC 3.1.3.18).
 GN Name=gph; ORFNames=BTZK2003;
 OS Bacillus cereus ZK.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=288681;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ZK;
 RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RT "Complete genome sequence of *Bacillus cereus* ZK.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; CP000001; AAU18252.1; -.
 KW Hydrolase.
 SQ SEQUENCE 221 AA; 25042 MW; C40390D934F50897 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 221;
 Best Local Similarity 58.3%; Pred. No. 72;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MYATEVLDDLGS 12
 || | : |||:
 Db 1 MYTTYLFDLDGT 12

RESULT 10

Q738Z3

ID Q738Z3 PRELIMINARY; PRT; 221 AA.
 AC Q738Z3;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hydrolase, haloacid dehalogenase-like family.
 GN OrderedLocusNames=BCE2250;
 OS *Bacillus cereus* (strain ATCC 10987).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 OX NCBI_TaxID=222523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14960714; DOI=10.1093/nar/gkh258;
 RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
 RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
 RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
 RT "The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic
 RT adaptations and a large plasmid related to *Bacillus anthracis* pXO1."
 RL Nucleic Acids Res. 32:977-988(2004).
 DR EMBL; AE017271; AAS41169.1; -.
 DR TIGR; BCE2250; -.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0008967; F:phosphoglycolate phosphatase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR005834; Dehal_like_hydro.
 DR InterPro; IPR006439; HAD_SF_A_v1.
 DR Pfam; PF00702; Hydrolase; 1.
 DR TIGRFAMs; TIGR01549; HAD-SF-IA-v1; 1.
 KW Complete proteome; Hydrolase.
 SQ SEQUENCE 221 AA; 25196 MW; E4359D20B5C50498 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 221;
 Best Local Similarity 58.3%; Pred. No. 72;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGS 12
||| : ||||:
Db 1 MYTTYLFDLDGT 12

RESULT 11

Q81R38

ID Q81R38 PRELIMINARY; PRT; 221 AA.
AC Q81R38; Q6HZB1; Q6KTA1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hydrolase, haloacid dehalogenase-like family.
GN OrderedLocusNames=BA2220, BAS2064, GBAA2220;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomason B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sterne;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE017031; AAP26098.1; -.
DR EMBL; AE017334; AAT31337.1; -.
DR EMBL; AE017225; AAT54378.1; -.
DR TIGR; BA2220; -.
DR TIGR; GBAA2220; -.

DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005834; Dehal_like_hydro.
DR Pfam; PF00702; Hydrolase; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 221 AA; 25147 MW; 4663727ADE72FA37 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 221;
Best Local Similarity 58.3%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDGS 12
||| : ||||:
Db 1 MYTTYLFDLDGT 12

RESULT 12

Q6HJE3

ID Q6HJE3 PRELIMINARY; PRT; 221 AA.
AC Q6HJE3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Phosphoglycolate phosphatase (EC 3.1.3.18).
GN Name=gph; OrderedLocusNames=BT9727_2004;
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=180856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE017355; AAT59749.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008967; F:phosphoglycolate phosphatase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR006439; HAD_SF_A_v1.
DR Pfam; PF00702; Hydrolase; 1.
DR TIGRFAMs; TIGR01549; HAD-SF-IA-v1; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 221 AA; 25173 MW; 5377232B9B69FB9C CRC64;

Query Match 60.0%; Score 39; DB 2; Length 221;
Best Local Similarity 58.3%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDGS 12
||| : ||||:
Db 1 MYTTYLFDLDGT 12

RESULT 13

Q8H9W4

ID Q8H9W4 PRELIMINARY; PRT; 315 AA.
 AC Q8H9W4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE ORF.49.
 GN Name=orf49;
 OS Pseudomonas aeruginosa phage PaP3.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.
 OX NCBI_TaxID=188350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PaP3;
 RA Hu F., Zhang K., Tan Y., Jin X., Zhu J., Huang J., Rao X., Shen X.,
 RA Hu X.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AY078382; AAL85522.1; -.
 SQ SEQUENCE 315 AA; 35754 MW; A4E4BDB15CDAC759 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 315;
 Best Local Similarity 63.6%; Pred. No. 1e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYATEVLDLDG 11
 ||| ||:| :|
 Db 3 MYAAEVIDREG 13

RESULT 14

Q88CQ1

ID Q88CQ1 PRELIMINARY; PRT; 362 AA.
 AC Q88CQ1;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=PP5129;
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weinell C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
 RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
 RA Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
 RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).

DR EMBL; AE016793; AAN70694.1; -.
DR TIGR; PP5129; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 362 AA; 40564 MW; 29B4A47B75BF14A3 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 362;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MYATEVLDDLGSK 13
|: ||| :||| :
Db 60 MFVTEVRELDGPR 72

RESULT 15

Q6CSK6

ID Q6CSK6 PRELIMINARY; PRT; 391 AA.
AC Q6CSK6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similarity.
GN ORFNames=KLLA0D00275g;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekaiia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; CR382124; CAH00179.1; -.
SQ SEQUENCE 391 AA; 41919 MW; C49252D91E05B5BF CRC64;

Query Match 60.0%; Score 39; DB 2; Length 391;
Best Local Similarity 72.7%; Pred. No. 1.3e+02;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YATEVLDLDGS 12
|:|||| | ||
Db 100 YSTEVLTLGS 110

Search completed: February 10, 2005, 15:57:36
Job time : 95.8451 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 52.3099 Seconds
(without alignments)
44.362 Million cell updates/sec

Title: US-10-067-484-9
Perfect score: 30
Sequence: 1 LLNNMR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	30	100.0	6	5	ABB81976	Abb81976 30 kDa ra
2	30	100.0	275	7	ADC33132	Adc33132 Human nov
3	30	100.0	358	3	AAB34358	Aab34358 Gene 6 hu
4	30	100.0	368	3	AAB34359	Aab34359 Human sec
5	30	100.0	423	8	ADP30146	Adp30146 Human sec
6	30	100.0	423	8	ADP30147	Adp30147 Human sec
7	30	100.0	450	6	ABR39840	Abr39840 Human SCA
8	30	100.0	450	8	ADN99870	Adn99870 Novel hum
9	30	100.0	459	6	AAE32103	Aae32103 Human cyt

10	30	100.0	459	8	ADN99321	Adn99321	Novel	hum
11	30	100.0	464	7	ADC31497	Adc31497	Human	nov
12	30	100.0	468	6	ABR39839	Abr39839	Human	SCA
13	30	100.0	482	4	ADM19812	Adm19812	Protein	e
14	30	100.0	494	8	ADP30148	Adp30148	Human	sec
15	30	100.0	494	8	ADP30149	Adp30149	Human	sec
16	30	100.0	495	8	ADP30150	Adp30150	Human	sec
17	30	100.0	495	8	ADP30151	Adp30151	Human	sec
18	30	100.0	511	4	ABG06408	Abg06408	Novel	hum
19	30	100.0	593	8	ADH13201	Adh13201	Human	mal
20	30	100.0	596	8	ADN99873	Adn99873	Novel	hum
21	30	100.0	596	8	ADN99872	Adn99872	Novel	hum
22	30	100.0	617	7	ADE31202	Ade31202	Human	dia
23	30	100.0	618	3	AAB56803	Aab56803	Human	pro
24	30	100.0	632	8	ADN99871	Adn99871	Novel	hum
25	30	100.0	1083	4	AAU39103	Aau39103	Mouse	mKC
26	30	100.0	1083	4	AAU39104	Aau39104	Mouse	mKC
27	30	100.0	1083	4	AAU39098	Aau39098	Mouse	pot
28	30	100.0	1083	7	ADE58573	Ade58573	Rat	Prote
29	30	100.0	1106	8	ADQ59403	Adq59403	Human	can
30	30	100.0	5002	4	ABB63723	Abb63723	Drosophil	
31	28	93.3	721	4	AAM25436	Aam25436	Human	pro
32	28	93.3	962	8	ADR09849	Adr09849	Human	pro
33	28	93.3	1099	4	AAU39096	Aau39096	Mouse	pot
34	28	93.3	1116	4	ABB11791	Abb11791	Human	K/C
35	28	93.3	1116	4	AAU39097	Aau39097	Human	pot
36	28	93.3	1135	4	AAU39094	Aau39094	Mouse	KCC
37	28	93.3	1135	4	AAU39093	Aau39093	Human	KCC
38	28	93.3	1150	4	AAU39095	Aau39095	Mouse	pot
39	28	93.3	1150	4	AAU39099	Aau39099	Human	pot
40	28	93.3	1155	4	AAM40061	Aam40061	Human	pol
41	28	93.3	1902	4	ABB65387	Abb65387	Drosophil	
42	28	93.3	3614	4	ABB62664	Abb62664	Drosophil	
43	27	90.0	46	4	AAM85893	Aam85893	Human	imm
44	27	90.0	172	6	ABU29589	Abu29589	Protein	e
45	27	90.0	322	5	ABP60872	Abp60872	Schizosac	

ALIGNMENTS

RESULT 1

ABB81976

ID ABB81976 standard; peptide; 6 AA.

XX

AC ABB81976;

XX

DT 25-NOV-2002 (first entry)

XX

DE 30 kDa ragweed pollen allergen tryptic peptide 9.

XX

KW Ragweed; pollen; allergen; Ambt 7; glycoprotein; antiallergic;
KW immunotherapy; disulphide protein.

XX

OS Ambrosia elatior.

XX

PN WO200263012-A2.

XX
 PD 15-AUG-2002.
 XX
 PF 04-FEB-2002; 2002WO-US003346.
 XX
 PR 05-FEB-2001; 2001US-0266686P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Buchanan BB, Del Val G, Frick OL;
 XX
 DR WPI; 2002-657539/70.
 XX
 PT New ragweed pollen allergens, useful in allergy testing and immunotherapy
 PT regimens, particularly for treating sensitivity to pollen or pollen
 PT allergy (e.g. anaphylaxis, or symptoms of hives or asthma) in a mammal,
 PT especially a human.
 XX
 PS Claim 1; Page 53; 70pp; English.
 XX
 CC The invention relates to an isolated pollen allergen purified from
 CC ragweed pollen, substantially free of any other pollen proteins, or a
 CC protein that is an antigenic fragment of a pollen allergen Ambt 7. The
 CC allergen is characterized by the following physiochemical and biological
 CC properties: (a) being contained in pollen extracts; (b) a glycoprotein;
 CC (c) a sulphhydryl group containing protein; (d) a molecular weight of
 CC about 30 kDa as determined by SDS-polyacrylamide gel electrophoresis; and
 CC (e) possessing allergen activity. The pollen allergen, or antigenic
 CC protein fragment of the pollen allergen Ambt 7, or composition is useful
 CC for treating sensitivity to pollen or pollen allergy in a mammal. This
 CC allergy includes anaphylaxis or atopy, which includes the symptoms of hay
 CC fever, asthma or hives. The allergen is also useful in allergy testing
 CC and immunotherapy regimens. Sequences ABB81968-978 represent tryptic
 CC peptide fragments of the 30 kDa ragweed complete pollen extract
 CC disulphide protein allergen
 XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 30; DB 5; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
 |||||
 Db 1 LLNNMR 6

RESULT 2
 ADC33132

ID ADC33132 standard; protein; 275 AA.
 XX
 AC ADC33132;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:3214.
 XX

KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 17.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029271-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 24-SEP-2002; 2002WO-US030474.
 XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX
 DR WPI; 2003-371981/35.
 DR N-PSDB; ADC32365.
 XX
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 PS Example 2; SEQ ID NO 3214; 1185pp; English.
 XX
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's

CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a human contig-
CC encoded polypeptide sequence used in an example of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 275 AA;

Query Match 100.0%; Score 30; DB 7; Length 275;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNNMR 6
|||
Db 76 LLNNMR 81

RESULT 3

AAB34358

ID AAB34358 standard; protein; 358 AA.

XX

AC AAB34358;

XX

DT 26-JAN-2001 (first entry)

XX

DE Gene 6 human secreted protein homologous amino acid sequence #119.

XX

KW Human; secreted protein; diagnosis; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active general; vulnerary;
KW gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW and antibacterial; gene therapy; detection; cancer; chromosome marker;
KW chromosome identification; neural disorder; immune disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; wound healing; infectious disease; preservative;
KW food additive.

XX

OS Mus musculus.

XX

PN WO200056883-A1.

XX

PD 28-SEP-2000.

XX

PF 16-MAR-2000; 2000WO-US006822.

XX

PR 23-MAR-1999; 99US-0126054P.

PR 10-DEC-1999; 99US-0169916P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Ruben SM, Komatsoulis G;

XX

DR WPI; 2000-587666/55.
 XX
 PT Human secreted proteins and gene sequences encoding them, useful for
 PT detecting, preventing, and treating disorders such as cancer,
 PT neurological disorders and immune system disorders.
 XX
 PS Disclosure; Page 391-392; 429pp; English.
 XX
 CC The polynucleotide sequences given in AAC59566 to AAC59614 encode the
 CC human secreted proteins given in AAB34299 to AAB34347. AAB34348 to
 CC AAB34437 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular
 CC active general; vulnerary; gastrointestinal; nephrotropic; antiinfective;
 CC gynaecological; and antibacterial. The polynucleotides can be used for
 CC the detection of various disorders such as cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The secreted proteins can be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wound healing, and infectious diseases. The proteins can also
 CC be used as a food additive or preservative to increase or decrease
 CC storage capabilities. AAC59557 to AAC59565 and AAB34298 represent
 CC sequences used in the exemplification of the present invention
 XX
 SQ Sequence 358 AA;

Query Match 100.0%; Score 30; DB 3; Length 358;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
 |||||
 Db 246 LLNNMR 251

RESULT 4

AAB34359

ID AAB34359 standard; protein; 368 AA.

XX

AC AAB34359;

XX

DT 26-JAN-2001 (first entry)

XX

DE Human secreted protein sequence encoded by gene 6 SEQ ID NO:120.

XX

KW Human; secreted protein; diagnosis; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active general; vulnerary;
 KW gastrointestinal; nephrotropic; antiinfective; gynaecological;
 KW and antibacterial; gene therapy; detection; cancer; chromosome marker;
 KW chromosome identification; neural disorder; immune disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW proliferative disorder; wound healing; infectious disease; preservative;
 KW food additive.

XX
 OS Homo sapiens.
 XX
 PN WO200056883-A1.
 XX
 PD 28-SEP-2000.
 XX
 PF 16-MAR-2000; 2000WO-US006822.
 XX
 PR 23-MAR-1999; 99US-0126054P.
 PR 10-DEC-1999; 99US-0169916P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX
 DR WPI; 2000-587666/55.
 XX
 PT Human secreted proteins and gene sequences encoding them, useful for
 PT detecting, preventing, and treating disorders such as cancer,
 PT neurological disorders and immune system disorders.
 XX
 PS Disclosure; Page 393-394; 429pp; English.
 XX
 CC The polynucleotide sequences given in AAC59566 to AAC59614 encode the
 CC human secreted proteins given in AAB34299 to AAB34347. AAB34348 to
 CC AAB34437 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular
 CC active general; vulnerary; gastrointestinal; nephrotropic; antiinfective;
 CC gynaecological; and antibacterial. The polynucleotides can be used for
 CC the detection of various disorders such as cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The secreted proteins can be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wound healing, and infectious diseases. The proteins can also
 CC be used as a food additive or preservative to increase or decrease
 CC storage capabilities. AAC59557 to AAC59565 and AAB34298 represent
 CC sequences used in the exemplification of the present invention
 XX
 SQ Sequence 368 AA;

Query Match 100.0%; Score 30; DB 3; Length 368;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
 |||||
 Db 256 LLNNMR 261

RESULT 5
 ADP30146
 ID ADP30146 standard; protein; 423 AA.

XX
AC ADP30146;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #913.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX

DR WPI; 2004-348438/32.

XX

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX

PS Claim 1; SEQ ID NO 2144; 428pp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.

XX

SQ Sequence 423 AA;

Query Match 100.0%; Score 30; DB 8; Length 423;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
|||||
Db 255 LLNNMR 260

RESULT 6

ADP30147

ID ADP30147 standard; protein; 423 AA.

XX

AC ADP30147;

XX

DT 12-AUG-2004 (first entry)

XX

DE Human secreted protein SEQ ID #914.

XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX

OS Homo sapiens.

XX

PN WO2004035732-A2.

XX

PD 29-APR-2004.

XX

PF 28-AUG-2003; 2003WO-US026780.

XX

PR 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX

DR WPI; 2004-348438/32.

XX

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX

PS Claim 1; SEQ ID NO 2145; 428pp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.

XX

SQ Sequence 423 AA;

Query Match 100.0%; Score 30; DB 8; Length 423;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNNMR 6
|||
Db 255 LLNNMR 260

RESULT 7

ABR39840

ID ABR39840 standard; protein; 450 AA.

XX

AC ABR39840;

XX

DT 11-AUG-2003 (first entry)

XX

DE Human SCAP polypeptide-Incyte Id. 3563232CD1.

XX

KW SCAP; structural and cytoskeleton-associated protein; nephrotropic;
KW cytostatic; antiarteriosclerotic; hepatotropic; virucide; antibacterial;
KW antihelminthic; cardiant; nootropic; neuroprotective; cerebroprotective;
KW anticonvulsant; gene therapy; transgenic; human.

XX

OS Homo sapiens.

XX

PN WO2003008625-A2.

XX

PD 30-JAN-2003.

XX

PF 19-JUL-2002; 2002WO-US022866.

XX

PR 20-JUL-2001; 2001US-0306810P.

PR 27-JUL-2001; 2001US-0308338P.

PR 07-AUG-2001; 2001US-0310980P.

PR 17-AUG-2001; 2001US-0313098P.

PR 31-AUG-2001; 2001US-0316796P.

PR 07-SEP-2001; 2001US-0317899P.

PR 14-SEP-2001; 2001US-0322183P.

PR 28-SEP-2001; 2001US-0326101P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Jones KA, Swarnakar A, Gorvad AE, Hafalia AJA, Warren BA;

PI Ison CH, Honchell CD, Nguyen DB, Barroso I, Das D, Lindquist EA;

PI Lee EA, Yue H, Forsythe IJ, Ramkumar J, Griffin JA, Li JX, Yang J;

PI Baughn MR, Borowsky ML, Thornton M, Yao MG, Walia NK, Burford N;

PI Lal PG, Gururajan R, Lee S, Bulloch SA, Becha SD, Richardson TW;

PI Elliott VS, Sprague WW, Tang YT, Azimzai Y, Lu Y, Zebarjadian Y;

XX
 DR WPI; 2003-239351/23.
 DR N-PSDB; ACC47270.
 XX
 PT New human structural and cytoskeleton-associated protein (SCAP), useful
 PT for diagnosing, treating and preventing diseases or conditions associated
 PT with aberrant SCAP expression, e.g. cancer, atherosclerosis or
 PT infections.
 XX
 PS Claim 1; Page 227-228; 267pp; English.
 XX
 CC The invention relates to novel human SCAP (structural and cytoskeleton-
 CC associated proteins and encoding polynucleotides. The SCAP polypeptides
 CC and polynucleotides are useful in diagnosing, treating and preventing
 CC diseases or conditions associated with aberrant expression of SCAP, such
 CC as cell motility disorders (e.g. ankylosing spondylitis), developmental
 CC disorders (e.g. renal tubular acidosis or dwarfism), cell proliferative
 CC disorders (e.g. cancer, arteriosclerosis, cirrhosis or hepatitis),
 CC infections (e.g. viral, bacterial or helminthic), heart and skeletal
 CC muscle disorders (e.g. muscular dystrophy or cardiomyopathy), and
 CC neurological disorders (e.g. Alzheimer's disease, Parkinson's disease,
 CC stroke, epilepsy or multiple sclerosis). These are also useful in
 CC assessing the effects of exogenous compounds on the expression of nucleic
 CC acid and amino acid sequences of SCAP. The SCAP or its fragments are
 CC useful in screening compounds for identifying modulators. The microarray
 CC is useful in monitoring or measuring protein-protein interactions, drug-
 CC target interactions, and gene expression profiles. Sequences ABR39805-841
 CC represent human SCAP polypeptides
 XX
 SQ Sequence 450 AA;

Query Match 100.0%; Score 30; DB 6; Length 450;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
 |||||
 Db 251 LLNNMR 256

RESULT 8

ADN99870

ID ADN99870 standard; protein; 450 AA.

XX

AC ADN99870;

XX

DT 29-JUL-2004 (first entry)

XX

DE Novel human protein sequence #686.

XX

KW anti-inflammatory; dermatological; neuroprotective; immunomodulator;

KW antibacterial; virucide; antipsoriatic; cytostatic; gene therapy;

KW vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;

KW early aging; hormonal imbalance; ischemic heart disease;

KW ulcerative colitis.

XX

OS Homo sapiens.

XX
PN WO2004038003-A2.
XX
PD 06-MAY-2004.
XX
PF 24-OCT-2003; 2003WO-US033947.
XX
PR 25-OCT-2002; 2002US-0421061P.
PR 25-OCT-2002; 2002US-0421080P.
PR 25-OCT-2002; 2002US-0421552P.
PR 25-OCT-2002; 2002US-0421614P.
PR 30-OCT-2002; 2002US-0422177P.
PR 30-OCT-2002; 2002US-0422178P.
PR 15-NOV-2002; 2002US-0426355P.
PR 15-NOV-2002; 2002US-0426384P.
PR 15-NOV-2002; 2002US-0426394P.
PR 15-NOV-2002; 2002US-0426430P.
PR 15-NOV-2002; 2002US-0426916P.
PR 27-NOV-2002; 2002US-0429224P.
PR 27-NOV-2002; 2002US-0429275P.
PR 27-NOV-2002; 2002US-0429302P.
PR 27-NOV-2002; 2002US-0429326P.
PR 27-NOV-2002; 2002US-0429651P.
PR 04-DEC-2002; 2002US-0430645P.
PR 04-DEC-2002; 2002US-0430651P.
PR 04-DEC-2002; 2002US-0430657P.
PR 04-DEC-2002; 2002US-0430663P.
PR 04-DEC-2002; 2002US-0430668P.
PR 04-DEC-2002; 2002US-0430684P.
PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 05-DEC-2002; 2002US-0431458P.
PR 12-DEC-2002; 2002US-0433251P.
PR 12-DEC-2002; 2002US-0433500P.
PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.

PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;

XX

DR WPI; 2004-365511/34.

DR N-PSDB; ADN99086.

XX

PT New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.

XX

PS Claim 14; SEQ ID NO 1470; 532pp; English.

XX

CC The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC protein of the invention.

XX

SQ Sequence 450 AA;

Query Match 100.0%; Score 30; DB 8; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
|||
Db 251 LLNNMR 256

RESULT 9

AAE32103

ID AAE32103 standard; protein; 459 AA.

XX

AC AAE32103;

XX

DT 24-MAR-2003 (first entry)

XX

DE Human cytoskeleton-associated protein, CSAP-1.
 XX
 KW Human; cytoskeleton-associated protein; CSAP-1; atherosclerosis; cancer;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1. .54
 FT /label= Signal_peptide
 FT Protein 55. .459
 FT /note= "Human mature CSAP-1"
 XX
 PN WO200279404-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 25-MAR-2002; 2002WO-US009288.
 XX
 PR 29-MAR-2001; 2001US-0280508P.
 PR 03-APR-2001; 2001US-0281323P.
 PR 13-APR-2001; 2001US-0283769P.
 PR 04-MAY-2001; 2001US-0288609P.
 PR 10-MAY-2001; 2001US-0290518P.
 PR 18-MAY-2001; 2001US-0291870P.
 PR 29-MAY-2001; 2001US-0294451P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Hafalia AJA, Tang TY, Yue H, Khan FA, Ison CH, Baughn MR;
 PI Warren BA, Duggan BM, Thangavelu K, Honchell CD, Azimzai Y;
 PI Elliott VS, Burford N, Ding L, Yue H, Becha S, Emerling BM;
 PI Richardson TW, Lee SY, Bandman O, Lal PG, Lee S, Gietzen KJ;
 PI Walia NK, Griffin JA, Lee EA, Swarnakar A, Ring HZ, Jones KA;
 XX
 DR WPI; 2003-092894/08.
 DR N-PSDB; AAD49590.
 XX
 PT New human cytoskeleton-associated proteins, useful for preparing a
 PT composition for diagnosing or treating a disease or condition associated
 PT with decreased expression or overexpression of functional CSAP e.g.,
 PT cancer.
 XX
 PS Claim 1; Page 148-149; 233pp; English.
 XX
 CC The invention relates to new human cytoskeleton-associated protein (CSAP)
 CC and its polynucleotide. The polypeptide is useful for preparing a
 CC composition for diagnosing or treating a disease or condition associated
 CC with decreased expression or overexpression of functional CSAP e.g.
 CC atherosclerosis or cancer. The present sequence is human CSAP-1 protein.
 CC The invention is useful in gene therapy
 XX
 SQ Sequence 459 AA;

Query Match 100.0%; Score 30; DB 6; Length 459;
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
|||||
Db 256 LLNNMR 261

RESULT 10

ADN99321

ID ADN99321 standard; protein; 459 AA.

XX

AC ADN99321;

XX

DT 29-JUL-2004 (first entry)

XX

DE Novel human protein sequence #137.

XX

KW anti-inflammatory; dermatological; neuroprotective; immunomodulator;

KW antibacterial; virucide; antipsoriatic; cytostatic; gene therapy;

KW vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;

KW early aging; hormonal imbalance; ischemic heart disease;

KW ulcerative colitis.

XX

OS Homo sapiens.

XX

PN WO2004038003-A2.

XX

PD 06-MAY-2004.

XX

PT 24-OCT-2003; 2003WO-US033947.

XX

PR 25-OCT-2002; 2002US-0421061P.

PR 25-OCT-2002; 2002US-0421080P.

PR 25-OCT-2002; 2002US-0421552P.

PR 25-OCT-2002; 2002US-0421614P.

PR 30-OCT-2002; 2002US-0422177P.

PR 30-OCT-2002; 2002US-0422178P.

PR 15-NOV-2002; 2002US-0426355P.

PR 15-NOV-2002; 2002US-0426384P.

PR 15-NOV-2002; 2002US-0426394P.

PR 15-NOV-2002; 2002US-0426430P.

PR 15-NOV-2002; 2002US-0426916P.

PR 27-NOV-2002; 2002US-0429224P.

PR 27-NOV-2002; 2002US-0429275P.

PR 27-NOV-2002; 2002US-0429302P.

PR 27-NOV-2002; 2002US-0429326P.

PR 27-NOV-2002; 2002US-0429651P.

PR 04-DEC-2002; 2002US-0430645P.

PR 04-DEC-2002; 2002US-0430651P.

PR 04-DEC-2002; 2002US-0430657P.

PR 04-DEC-2002; 2002US-0430663P.

PR 04-DEC-2002; 2002US-0430668P.

PR 04-DEC-2002; 2002US-0430684P.

PR 05-DEC-2002; 2002US-0430937P.

PR 05-DEC-2002; 2002US-0430965P.

PR 05-DEC-2002; 2002US-0431458P.

PR 12-DEC-2002; 2002US-0433251P.

PR 12-DEC-2002; 2002US-0433500P.

PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;

XX

DR WPI; 2004-365511/34.

DR N-PSDB; ADN98537.

XX

PT New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.

XX

PS Claim 14; SEQ ID NO 921; 532pp; English.

XX

CC The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,

CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC protein of the invention.
XX
SQ Sequence 459 AA;

Query Match 100.0%; Score 30; DB 8; Length 459;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
|||
Db 256 LLNNMR 261

RESULT 11

ADC31497

ID ADC31497 standard; protein; 464 AA.

XX

AC ADC31497;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human novel polypeptide sequence, SEQ ID NO:1579.

XX

KW Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

KW gene therapy; chromosome 17.

XX

OS Homo sapiens.

XX

PN WO2003029271-A2.

XX

PD 10-APR-2003.

XX

PF 24-SEP-2002; 2002WO-US030474.

XX

PR 24-SEP-2001; 2001US-0324631P.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

PI Haley-Vicente D, Drmanac RT;

XX

DR WPI; 2003-371981/35.

DR N-PSDB; ADC30526.

XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or

PT treating conditions such as neurodegenerative diseases, anemias, platelet

PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

PT cancer.

XX
PS Claim 20; SEQ ID NO 1579; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 464 AA;

Query Match 100.0%; Score 30; DB 7; Length 464;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
| | | | |
Db 258 LLNNMR 263

RESULT 12
ABR39839
ID ABR39839 standard; protein; 468 AA.
XX
AC ABR39839;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human SCAP polypeptide-Incyte Id. 7502011CD1.
XX

KW SCAP; structural and cytoskeleton-associated protein; nephrotropic;
 KW cytostatic; antiarteriosclerotic; hepatotropic; virucide; antibacterial;
 KW antihelminthic; cardiant; nootropic; neuroprotective; cerebroprotective;
 KW anticonvulsant; gene therapy; transgenic; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003008625-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 19-JUL-2002; 2002WO-US022866.
 XX
 PR 20-JUL-2001; 2001US-0306810P.
 PR 27-JUL-2001; 2001US-0308338P.
 PR 07-AUG-2001; 2001US-0310980P.
 PR 17-AUG-2001; 2001US-0313098P.
 PR 31-AUG-2001; 2001US-0316796P.
 PR 07-SEP-2001; 2001US-0317899P.
 PR 14-SEP-2001; 2001US-0322183P.
 PR 28-SEP-2001; 2001US-0326101P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Jones KA, Swarnakar A, Gorvad AE, Hafalia AJA, Warren BA;
 PI Ison CH, Honchell CD, Nguyen DB, Barroso I, Das D, Lindquist EA;
 PI Lee EA, Yue H, Forsythe IJ, Ramkumar J, Griffin JA, Li JX, Yang J;
 PI Baughn MR, Borowsky ML, Thornton M, Yao MG, Walia NK, Burford N;
 PI Lal PG, Gururajan R, Lee S, Bulloch SA, Becha SD, Richardson TW;
 PI Elliott VS, Sprague WW, Tang YT, Azimzai Y, Lu Y, Zebarjadian Y;
 XX
 DR WPI; 2003-239351/23.
 DR N-PSDB; ACC47269.
 XX
 PT New human structural and cytoskeleton-associated protein (SCAP), useful
 PT for diagnosing, treating and preventing diseases or conditions associated
 PT with aberrant SCAP expression, e.g. cancer, atherosclerosis or
 PT infections.
 XX
 PS Claim 1; Page 225-226; 267pp; English.
 XX
 CC The invention relates to novel human SCAP (structural and cytoskeleton-
 CC associated proteins and encoding polynucleotides. The SCAP polypeptides
 CC and polynucleotides are useful in diagnosing, treating and preventing
 CC diseases or conditions associated with aberrant expression of SCAP, such
 CC as cell motility disorders (e.g. ankylosing spondylitis), developmental
 CC disorders (e.g. renal tubular acidosis or dwarfism), cell proliferative
 CC disorders (e.g. cancer, arteriosclerosis, cirrhosis or hepatitis),
 CC infections (e.g. viral, bacterial or helminthic), heart and skeletal
 CC muscle disorders (e.g. muscular dystrophy or cardiomyopathy), and
 CC neurological disorders (e.g. Alzheimer's disease, Parkinson's disease,
 CC stroke, epilepsy or multiple sclerosis). These are also useful in
 CC assessing the effects of exogenous compounds on the expression of nucleic
 CC acid and amino acid sequences of SCAP. The SCAP or its fragments are
 CC useful in screening compounds for identifying modulators. The microarray
 CC is useful in monitoring or measuring protein-protein interactions, drug-
 CC target interactions, and gene expression profiles. Sequences ABR39805-841

CC represent human SCAP polypeptides
XX
SQ Sequence 468 AA;

Query Match 100.0%; Score 30; DB 6; Length 468;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNNMR 6
|||
Db 255 LLNNMR 260

RESULT 13

ADM19812

ID ADM19812 standard; protein; 482 AA.

XX

AC ADM19812;

XX

DT 20-MAY-2004 (first entry)

XX

DE Protein encoded by novel human channel/transporter gene #130.

XX

KW immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW gene therapy; channel/transporter protein; rheumatoid arthritis;
KW neoplasm; cardiac arrest; cerebrovascular disorder; cerebral ischemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease;
KW ocular disorder; corneal infection; wound healing;
KW epithelial cell proliferation; skin aging; sunburn; transplantation;
KW chemotaxis; food additive.

XX

OS Homo sapiens.

XX

PN WO200154472-A2.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001307.

XX

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 03-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Barash SC, Ruben SM;

XX

DR WPI; 2001-476159/51.

DR N-PSDB; ADM19333.

XX

PT Isolated nucleic acid molecule encoding a channel/transporter protein is
PT used in preventing, treating or ameliorating a medical condition.

XX

PS Claim 11; SEQ ID NO 619; 809pp; English.

XX

CC The invention relates to an isolated nucleic acid molecule encoding a
CC channel/transporter protein or sequences at least 95% identical to a
CC these. The nucleic acids and proteins encoded by them are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. The antibodies to the proteins can also be used
CC in alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. This sequence corresponds to a protein of the
CC invention.

XX

SQ Sequence 482 AA;

Query Match 100.0%; Score 30; DB 4; Length 482;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6

|||||

Db 276 LLNNMR 281

RESULT 14

ADP30148

ID ADP30148 standard; protein; 494 AA.

XX

AC ADP30148;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #915.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX

DR WPI; 2004-348438/32.

XX

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX

PS Claim 1; SEQ ID NO 2146; 428pp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.

XX

SQ Sequence 494 AA;

Query Match 100.0%; Score 30; DB 8; Length 494;

Best Local Similarity 100.0%; Pred. No. 5.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
|||||
Db 255 LLNNMR 260

RESULT 15

ADP30149

ID ADP30149 standard; protein; 494 AA.

XX

AC ADP30149;

XX

DT 12-AUG-2004 (first entry)

XX

DE Human secreted protein SEQ ID #916.

XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.

XX

OS Homo sapiens.

XX

PN WO2004035732-A2.

XX

PD 29-APR-2004.

XX

PF 28-AUG-2003; 2003WO-US026780.

XX

PR 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX

DR WPI; 2004-348438/32.

XX

PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX

PS Claim 1; SEQ ID NO 2147; 428pp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.

XX

SQ Sequence 494 AA;

Query Match 100.0%; Score 30; DB 8; Length 494;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
|||
Db 255 LLNNMR 260

Search completed: February 10, 2005, 15:48:46
Job time : 54.3099 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 13.4366 Seconds
(without alignments)
33.334 Million cell updates/sec

Title: US-10-067-484-9
Perfect score: 30
Sequence: 1 LLNNMR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	30	100.0	22	1	US-07-934-656A-21	Sequence 21, Appl
2	30	100.0	593	4	US-09-538-092-919	Sequence 919, App
3	28	93.3	223	4	US-09-270-767-38461	Sequence 38461, A
4	28	93.3	223	4	US-09-270-767-53678	Sequence 53678, A
5	28	93.3	1116	4	US-09-949-016-10408	Sequence 10408, A
6	27	90.0	129	4	US-09-270-767-32999	Sequence 32999, A
7	27	90.0	129	4	US-09-270-767-48216	Sequence 48216, A
8	27	90.0	302	4	US-09-248-796A-24789	Sequence 24789, A
9	27	90.0	358	4	US-09-107-532A-4677	Sequence 4677, Ap
10	27	90.0	475	4	US-09-248-796A-15218	Sequence 15218, A
11	27	90.0	637	4	US-09-543-681A-6512	Sequence 6512, Ap
12	27	90.0	726	4	US-09-489-039A-7465	Sequence 7465, Ap
13	27	90.0	814	4	US-09-328-352-4373	Sequence 4373, Ap
14	27	90.0	1275	4	US-09-749-340-6	Sequence 6, Appli
15	26	86.7	65	4	US-09-107-532A-4450	Sequence 4450, Ap
16	26	86.7	161	4	US-09-248-796A-15514	Sequence 15514, A
17	26	86.7	163	4	US-09-489-039A-12148	Sequence 12148, A
18	26	86.7	217	4	US-09-328-352-7768	Sequence 7768, Ap
19	26	86.7	278	4	US-09-252-991A-21606	Sequence 21606, A
20	26	86.7	299	4	US-09-270-767-61405	Sequence 61405, A
21	26	86.7	337	3	US-08-448-722A-2	Sequence 2, Appli
22	26	86.7	337	3	US-08-189-309B-2	Sequence 2, Appli
23	26	86.7	436	3	US-08-486-099-94	Sequence 94, Appl
24	26	86.7	436	3	US-08-360-107A-104	Sequence 104, App
25	26	86.7	436	3	US-08-484-223B-94	Sequence 94, Appl
26	26	86.7	436	3	US-08-919-597-94	Sequence 94, Appl
27	26	86.7	436	3	US-08-475-668A-94	Sequence 94, Appl
28	26	86.7	436	3	US-08-485-551A-94	Sequence 94, Appl
29	26	86.7	436	3	US-08-471-913A-94	Sequence 94, Appl
30	26	86.7	436	3	US-08-485-264A-94	Sequence 94, Appl
31	26	86.7	436	3	US-08-474-349A-94	Sequence 94, Appl
32	26	86.7	436	4	US-08-255-208A-30	Sequence 30, Appl
33	26	86.7	436	4	US-08-470-896-94	Sequence 94, Appl
34	26	86.7	436	4	US-08-485-546A-94	Sequence 94, Appl
35	26	86.7	436	4	US-08-487-266A-94	Sequence 94, Appl
36	26	86.7	444	4	US-09-350-841A-1585	Sequence 1585, Ap
37	26	86.7	463	4	US-09-345-236B-2	Sequence 2, Appli
38	26	86.7	498	4	US-09-543-681A-7590	Sequence 7590, Ap
39	26	86.7	513	5	PCT-US91-02714-26	Sequence 26, Appl
40	26	86.7	553	2	US-08-663-566A-13	Sequence 13, Appl
41	26	86.7	553	2	US-08-484-575A-14	Sequence 14, Appl
42	26	86.7	553	2	US-08-023-610-13	Sequence 13, Appl
43	26	86.7	553	2	US-08-288-065A-13	Sequence 13, Appl
44	26	86.7	553	2	US-08-362-240A-13	Sequence 13, Appl
45	26	86.7	553	3	US-08-477-459-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-07-934-656A-21

; Sequence 21, Application US/07934656A

; Patent No. 5500347

; GENERAL INFORMATION:

; APPLICANT: MOLL, Roland

; APPLICANT: FRANKE, Werner W.

; TITLE OF INVENTION: PROCESS FOR THE PURIFICATION OF

; TITLE OF INVENTION: CYTOKERATIN 20 AND ITS USE FOR THE PRODUCTION OF

; TITLE OF INVENTION: ANTIBODIES

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram

; STREET: 655 Fifteenth Street N.W. Suite 330

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005-5701

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/934,656A

; FILING DATE: 27-JAN-1993

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: DE P 40 23 945.4

; FILING DATE: 27-JUL-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Murray, Robert B.

; REGISTRATION NUMBER: 22,980

; REFERENCE/DOCKET NUMBER: P564-3003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)638-5000

; TELEFAX: (202)638-4810

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-07-934-656A-21

Query Match 100.0%; Score 30; DB 1; Length 22;

Best Local Similarity 100.0%; Pred. No. 6.5;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6

|||||

Db 16 LLNNMR 21

RESULT 2

US-09-538-092-919

; Sequence 919, Application US/09538092

```

; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 919
;   LENGTH: 593
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number P13645
US-09-538-092-919

```

```

Query Match          100.0%;  Score 30;  DB 4;  Length 593;
Best Local Similarity 100.0%;  Pred. No. 1.7e+02;
Matches      6;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy          1 LLNNMR 6
             |||||
Db          317 LLNNMR 322

```

RESULT 3

```

US-09-270-767-38461
; Sequence 38461, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38461
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-38461

```

```

Query Match          93.3%;  Score 28;  DB 4;  Length 223;
Best Local Similarity 83.3%;  Pred. No. 1.7e+02;
Matches      5;  Conservative    1;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy          1 LLNNMR 6
             |:|||

```

Db 2 LINNMR 7

RESULT 4

US-09-270-767-53678
; Sequence 53678, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53678
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-53678

Query Match 93.3%; Score 28; DB 4; Length 223;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
|:||||
Db 2 LINNMR 7

RESULT 5

US-09-949-016-10408
; Sequence 10408, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES
THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10408
; LENGTH: 1116
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10408

Query Match 93.3%; Score 28; DB 4; Length 1116;

Best Local Similarity 83.3%; Pred. No. 8.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
:|||||
Db 280 MLNNMR 285

RESULT 6

US-09-270-767-32999
; Sequence 32999, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32999
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32999

Query Match 90.0%; Score 27; DB 4; Length 129;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
:|||||
Db 16 LLNNLR 21

RESULT 7

US-09-270-767-48216
; Sequence 48216, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48216
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-48216

Query Match 90.0%; Score 27; DB 4; Length 129;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
| | | | : |
Db 16 LLNNLR 21

RESULT 8

US-09-248-796A-24789
; Sequence 24789, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA
ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24789
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24789

Query Match 90.0%; Score 27; DB 4; Length 302;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6
: | | | |
Db 169 VLNNMR 174

RESULT 9

US-09-107-532A-4677
; Sequence 4677, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND
THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC